57

CU/E-H03255-57-P014903

AMERICAN
OURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

T. H. HILDEBRANDT
UNIVERSITY OF MICHIGAN

H. WEYL
THE INSTITUTE FOR ADVANCED STUDY

F. D. MURNAGHAN
E JOHNS HOPKINS UNIVERSITY

R. L. WILDER
UNIVERSITY OF MICHIGAN

O. ZARISKI
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

C. R. ADAMS
R. D. JAMES
SAUNDERS MacLANE
GABOR SZEGŐ
LEO ZIPPIN

G. BIRKHOFF
N. DUNFORD
E. P. LANE
D. MONTGOMERY
J. L. SYNGE

R. P. AGNEW
C. CHEVALLEY
G. A. HEDLUND
S. B. MYERS
N. E. STEENROD

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

510.5-
002

VOLUME LXIV

1942

CALCUTTA UNIVERSITY CENTRAL LIBRARY

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

914205

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

INDEX

PAGE

ronszajn, Nathan and Weinstein, Alexander. On the unified theory of eigenvalues of plates and membranes, ... 623

Baer, Reinhold. Homogeneity of projective planes, ... 137

Banerjee, D. P.. The expansion of an arbitrary function in a series of toroidal functions of the second kind, ... 72

ankier, J. D. and Leighton, Walter. Numerical continued fractions, 653

Beaumont, Ross A. Projections of non-abelian groups upon abelian groups containing elements of infinite order, ... 115

Begle, Edward G. Locally connected spaces and generalized manifolds, 553

Bell, P. O. The parametric osculating quadrics of a family of curves on a surface, ... 343

Bers, Lipman. On bounded analytic functions of two complex variables in certain domains with distinguished boundary surface, . . 514

Bourgin, D. G. Some properties of Banach spaces, ... 597

Brauer, Alfred. On a problem of partitions, ... 299

Brauer, Richard. On groups whose order contains a prime number to the first power I, ... 401

Brauer, Richard. On groups whose order contains a prime number to the first power II, ... 421

Bruck, Richard H. and Wade, T. L. Bisymmetric tensor algebra. Part I, ... 725

Bruck, Richard H. and Wade, T. L. Bisymmetric tensor algebra. Part II, ... 734

Chojnacki, Chaim. A contribution to the four color problem, . . 36

Clifford, A. H. Matrix representations of completely simple semigroups, 327

Coburn, N. Semi-analytic unitary subspaces of unitary space, . . 714

Comenetz, George. The limit of the ratio of arc to chord, . . . 695

Davenport, H. On sums of positive integral k-th powers, . . . 189

Davenport, H. On Waring's problem for fifth and sixth powers, . 199

PAGE

DAVIDS, NORMAN. Minimal surfaces spanning closed manifolds and having prescribed topological position, ... 348

EILENBERG, SAMUEL and WILDER, R. L. Uniform local connectedness and contractibility, ... 613

EVERETT, C. J., JR. An extension theory for rings, ... 363

GARABEDIAN, H. L. A class of linear integral transformations, . . 208

GASKELL, R. E. A problem in heat conduction and an expansion theorem, 447

HALMOS, PAUL R. Square roots of measure preserving transformations, 153

HARTMAN, PHILIP and WINTNER, AUREL. On the infinitesimal generators of integral convolutions, ... 273

HOCHSCHILD, G. Semi-simple algebras and generalized derivations, . 677

INGHAM, A. E. On two conjectures in the theory of numbers, . . 313

JAMES, R. D. and WEYL, H. Elementary note on prime number problems of Vinogradoff's type, ... 539

KAPLAN, WILFRED. The structure of a curve-family on a surface in the neighborhood of an isolated singularity, ... 1

LEHMER, D. H. Properties of the coefficients of the modular invariant $J(\tau)$, ... 488

LEIGHTON, WALTER. See Bankier and Leighton, page 653.

LEVI, HOWARD. Composite polynomials with coefficients in an arbitrary field of characteristic zero, ... 389

LEWIS, PAUL E. Characters of abelian groups, ... 81

LONSETH, A. T. The problem of Plateau in hyperbolic space, . . 229

MACDUFFEE, C. C. Products and norms of ideals, ... 646

MILLER, D. D. Extension and reduction theorems for certain types of continuous transformations, ... 215

MUHLY, H. T. Valuations and infinitely near algebraic loci, . . 457

OLDENBURGER, RUFUS. Repeated linear factors of forms, . . . 592

OLLMANN, LOYAL F. Solution of a problem of Ayres, . . . 61

PALL, GORDON. Quaternions and sums of three squares, . . . 503

PERLIN, IRWIN E. Indefinitely differentiable functions in several real variables, ... 441

RADEMACHER, H. A. Correction, ... 456

REISSNER, ERIC. Note on the expressions for the strains in a bent, thin shell, ... 768

PAGE

ROBINSON, CHAS. V. Spherical theorems of Helly type and congruence indices of spherical caps, ... 260

ROBINSON, G. DE B. Note on a paper by R. H. Bruck and T. L. Wade, . 753

ROSENBLOOM, PAUL C. Post algebras. I. Postulates and general theory, 167

SALEM, R. On sets of multiplicity for trigonometrical series, . . 531

SZÁSZ, OTTO. On convergence and summability of trigonometric series, 575

THOMAS, T. Y. Qualitative analysis of the flow of fluids in pipes, . 754

THRALL, R. M. On symmetrized Kronecker powers and the structure of the free Lie ring, ... 371

WADE, T. L. See Bruck and Wade, page 725.

WADE, T. L. See Bruck and Wade, page 734.

WEINSTEIN, ALEXANDER. See Aronszajn and Weinstein, page 623.

WEISNER, LOUIS. Polynomials whose roots lie in a sector, ... 55

WEYL, H. See James and Weyl, page 539.

WILDER, R. L. See Eilenberg and Wilder, page 613.

WINTNER, AUREL. On a statistics of the Ramanujan sums, . . . 106

WINTNER, AUREL. On the prime number theorem, ... 320

WINTNER, AUREL. On the Ω-matrices of Toeplitz, ... 669

WINTNER, AUREL. See Hartman and Wintner, page 273.

CALCUTTA UNIVERSITY CENTRAL LIBRARY

THE STRUCTURE OF A CURVE-FAMILY ON A SURFACE IN THE NEIGHBORHOOD OF AN ISOLATED SINGULARITY.*

By WILFRED KAPLAN.

Introduction. 1. The two problems. In this paper two problems will be considered: (A) to determine and classify topologically the possible structures of a curve-family on the surface of a sphere with precisely two singularities; (B) to determine the possible structures of a curve-family on any surface in the neighborhood of an isolated singularity. It will be shown that problem (B) is included in problem (A).

2. **Regularity. Families filling the plane.** The curve-families considered will be assumed *regular:* i. e., locally homeomorphic with parallel lines, except at singularities. Two previous papers by the author[1] have considered the possible structures of a regular curve-family filling the plane without singularity (topologically the same as a family on the sphere with one singularity). A complete topological classification was obtained, as follows: each family was shown to consist of open curves tending to infinity in both directions; each family was associated in a one-to-one manner with a set of non-intersecting chords (or *chordal system*) on a circle, the relative positions of curves of the family being reflected in the corresponding chords; it was shown that two curve-families are homeomorphic if and only if the corresponding chordal systems have the same arrangement; the exact specifications of the types of chordal systems appearing were obtained.

Many of the results of the present paper will be obtained by referring back to results of the two previous papers.

3. **The three cases.** For the problem (A) of families filling the sphere with the exception of two singular points N and S we consider three cases: (a) there is a curve C of the family joining N and S; (b) there is no such curve, nor is there any closed curve in the family; (c) there is a closed curve in the family. These cases exhaust all possibilities.

---

* Received February 12, 1941; Revised August 26, 1941; Presented to the American Mathematical Society April 26, 1940.

[1] See Kaplan [5] and [6]. (Numbers in brackets refer to the bibliography at the end.) These papers will be referred to by the symbols I and II respectively and a theorem or section of one of them will be indicated by prefacing I or II to the number of the theorem or section.

4. **Procedure in case (a).** In case (a) we cut open the sphere minus *N* and *S* along the curve *C* and map on a strip region of the plane whose two boundary curves correspond to *C*. The possibilities in a strip region in the plane were determined in the previous papers; hence this case is settled immediately. In Figure 1 a curve-family of case (a) is illustrated.

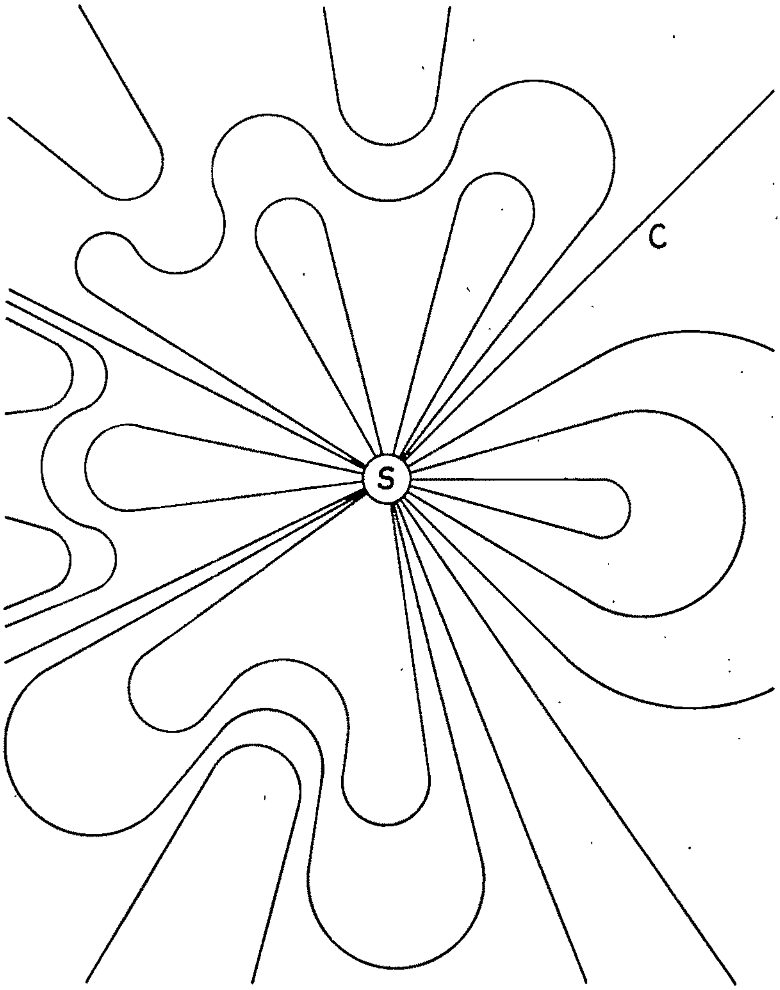


Fig. 1. Illustrating a curve-family on the sphere with two singularities *N* and *S*, case (a): there is a curve *C* of the family joining *N* and *S*. The sphere is here drawn stereographically as a plane, with *N* at infinity.

5. **Procedure in case (b).** In case (b) the singularity at *N* or *S* is locally the same as in case (a). The family can be decomposed into three

distinct sub-families, which we call $B(N)$, J, and $B(S)$. $B(N)$ consists of a set of bounding curves which we call $Fl(N)$, the "flower at N," plus curves inside the petals of the flower. This is illustrated in Figure 2. $B(S)$ has the same type of structure. J consists of the "joining" curves, open curves whose limiting set in one direction is in $B(N)$ and in the other direction in

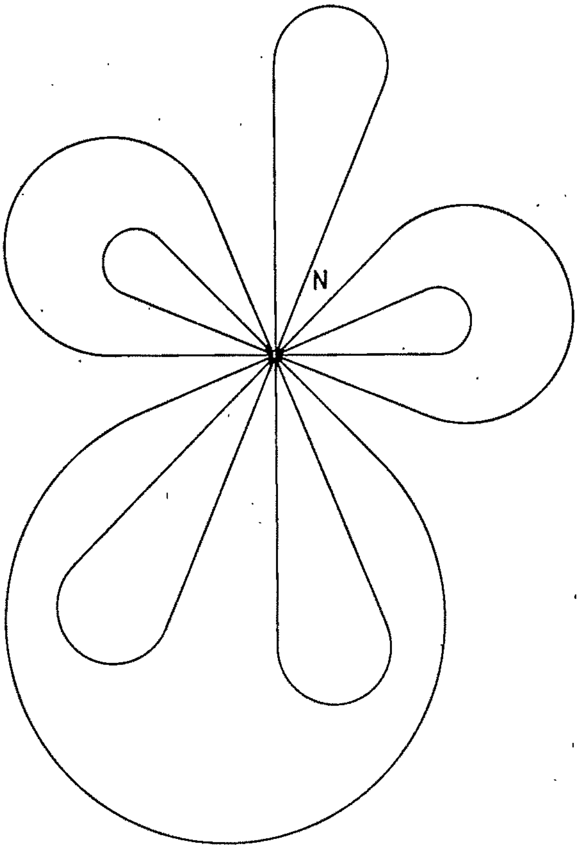


Fig. 2. Illustrating a 'flower at N' plus curves inside the 'petals.'

$B(S)$. If $B(N)$ and $B(S)$ were both shrunk to points, then J would appear as a family like the parallels of longitude on the globe. The curves of J may approach $B(N)$ in one of two ways: all curves may spiral around $B(N)$ infinitely often; or each curve may directly find its way to N without spiraling. In the latter case the picture at N is a special case of case (a). The two ways in which J and $B(N)$ can fit together are illustrated in Figures 3 and 4. The same possibilities can occur at S, and by considering the various combinations of structures at N and S we obtain all the possible structures in the large. It is to be noted that J must spiral around either $B(N)$ or $B(S)$, or else we are in case (a) again.

6. **Procedure in case (c).** In case (c) the closed curves are placed like latitudinal circles and decompose the family into annular regions. The possible families in an annulus will be analyzed in detail. Such families are built up from two fundamental types: the type of spirals which are asymptotic to the two bounding closed curves in opposite directions; the type of spirals which are asymptotic to the two bounding closed curves in the same direction. See Figures 5a and 5b.

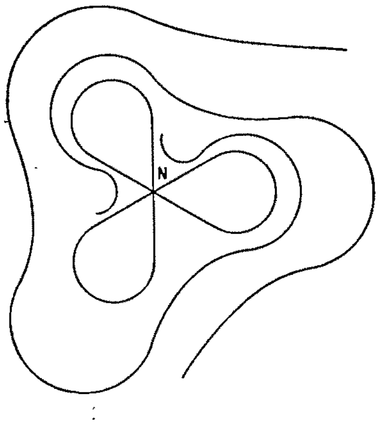


Fig. 3. Spirals around a flower.

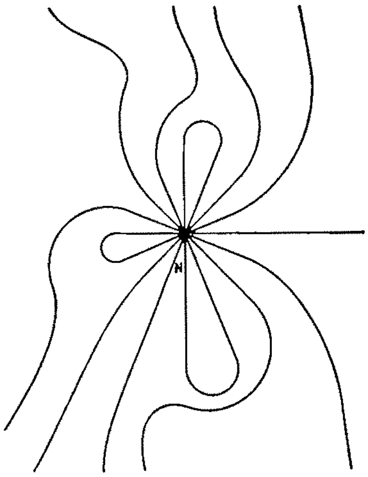


Fig. 4. Curves of J go directly to N.

At N or S there may be a last closed curve D, in which case the structure of the family near the singularity may be that of spirals leading from D to the singular point; if there is no last closed curve at N or S, the set of closed curves has as limit either a flower at N or S or the point N or S itself.

7. **Methods used.** The basic tool is that of cutting open and identifying.[2] In particular the analysis of the annular families (which occur in a certain sense in case (b) besides in case (c)) will depend on the process of reducing a family of lines filling a rectangle to a family of curves filling an annulus by identifying opposite points of a pair of sides.

---

[2] See H. Seifert and W. Threlfall, *Lehrbuch der Topologie* (Leipzig and Berlin, 1934), pp. 3-13, 31-35.

This method will be applied in case (b) to study the situation illustrated in Figure 3. It is to be noted that the flower there acts like a simple closed curve as far as the spirals are concerned. In the general case, when there may be infinitely many petals, this is not obvious. It will be established that it is possible to 'explode' the flower in such a way that the exterior region plus the curves C_n of the flower are mapped homeomorphically on the exterior of a circle plus a set of open arcs C'_n on the circumference. This homeomorphism can be so chosen that the image family is regular at every point of the circle.

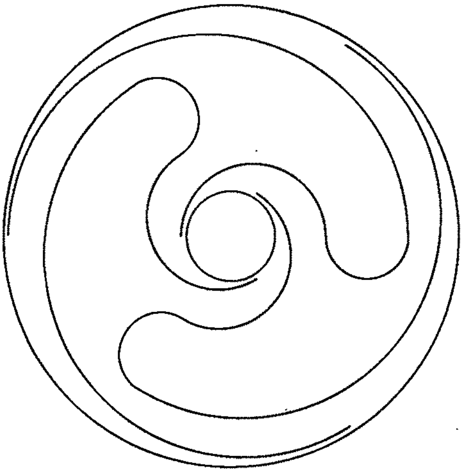


Fig. 5a.

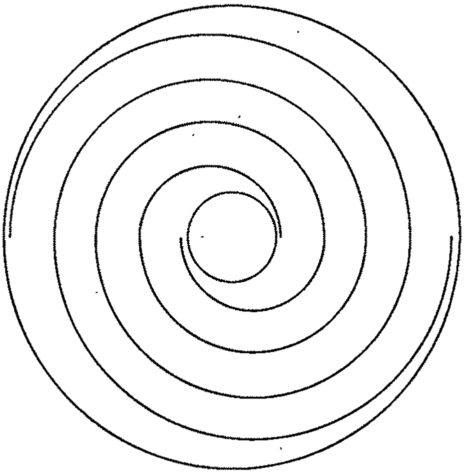


Fig. 5b.

In other words, the family illustrated in Figure 3 can be imbedded in a regular annular family. The proof of this fact rests on a cutting open of the family to obtain a sub-family of a regular family filling the plane and on taking advantage of the results of I and II.

8. **Problem B.** For the problem (B) we use the following device: We choose a simply-connected neighborhood $U(p)$ of the isolated singular point p, so that the family is regular in $U(p) - p$. We then map $U(p) - p$ homeomorphically on the sphere minus N and S. Our given family in $U(p) - p$ then becomes a family on the sphere with singularities N and S. The results of problem (A) are then applied to describe all possible structures in $U(p) - p$.

PART A. Families on the Sphere with Two Singularities.

1. General properties.

1. 1. **Curve-families on the sphere.** By an *open curve* will be meant a homeomorphic image of an open interval. By a *closed curve* will be meant a homeomorphic image of a circle. By an *arc* will be meant a homeomorphic image of a closed interval. By a *half-open curve* will be meant a homeomorphic image of a half-open interval. The word *curve* will denote any one of the preceding four.

The sphere $x^2 + y^2 + (z - 1)^2 = 1$ in xyz-space will be denoted by K. The xy-plane will be denoted by π.

A family F of curves $\{C\}$ will be said to *fill* a subset R of K or π if each curve C is in R and each point of R is on one and only one curve C. Let R_0 be a subset of K such that each curve C of F intersects R_0 in a set $R_0 \cdot C$ each of whose components is a curve. Then $F[R_0]$ will denote the curve-family filling $R_0 \cdot R$ whose curves are the different components of all the sets $R_0 \cdot C$, for C in F.

If G is any collection of curves on K or π, G will also denote the point-set union of the curves of G; the context will always make the meaning unambiguous.

If the family F_1 fills R_1 and the family F_2 fills R_2, then F_1 and F_2 will be termed *homeomorphic* if there is a homeomorphism of R_1 on R_2 such that the image of each curve of F_1 is a curve of F_2.

1. 2. **Regularity.** Throughout this section F will denote a curve-family filling a region R of K or π. R^i will denote the interior of R, $\bar{R}$ its closure. We shall assume $R^i \neq 0$. Let F_0 denote the family of lines $y =$ constant of the rectangle $R_0 : |x| \leqq 1, |y| \leqq 1$. F will be said to be *regular* at a point p of R if there is a set $U(p)$ such that p is interior to $U(p)$ and such that $F[U(p)]$ is homeomorphic to F_0. (Thus in particular F_0 is regular at each point, interior or boundary, of the rectangle.) F is termed *regular in* R if it is regular at each point of R. A cross-section of F (through the point r) shall mean an arc pq (to which r is interior) which lies in a subset R_1 of R which is open relative to R and is such that $F[R_1]$ is defined and each curve of $F[R_1]$ meets pq at most once.

An *r-neighborhood* of a point p of R shall mean a set $U(p)$ which contains p and is open relative to R and whose closure $\overline{U(p)}$ lies in R and can be mapped homeomorphically on the closed rectangle $|x| \leqq 1$, $|y| \leqq 1$ in such a way that $F[\overline{U(p)}]$ becomes the family F_0, while the inverse images of the lines $|x| = 1$ are cross-sections of F.

By Theorem I. 5 and the above definitions we have

(i) *If F is regular in R, then every point p of R has an arbitrarily small r-neighborhood $U(p)$, and there is a cross-section pq with p as end-point. If further $p \subset R^i$, then there is a cross-section through p.*

(ii) *Let $R = R_1 + R_2$, where $F[R_1]$ and $F[R_2]$ are both defined, let p be an interior point of R, and let $F[R_1]$ and $F[R_2]$ be regular at p. Then F is regular at p.*

For example, take F as a curve-family filling the interior R of a circle. Let one curve of F be a diameter minus its end-points and let R_1 and R_2 be the interiors of the corresponding semi-circles.

1.3. The *s*-families.[3] By a homeomorphism $y' = f(y)$ defined for $0 \leqq y \leqq 2$, and with $f(0) < f(1) = 1 < f(2)$, points $(0, y)$ on the line $x = 0$, $0 \leqq y \leqq 2$ can be identified with points $(1, y')$ on the line $x = 1$. With this identification made, the rectangle $0 \leqq x \leqq 1$, $0 \leqq y \leqq 2$ plus the points $(1, y')$ for $f(0) \leqq y' \leqq f(2)$ becomes homeomorphic to a region G of π. Further, the family of lines $y =$ constant of the rectangle becomes a curve-family F_1 filling G.

Any curve-family homeomorphic to a family $F_1[G^i]$ will be termed an *open s-family.* Any curve-family homeomorphic to a family $F_1[G^i \cdot G^*]$, where G^* is the set of images of the points (x, y) of the rectangle with $y \geqq 1$, will be termed a *half-open s-family.* By a *closed s-family* will be meant a curve-family homeomorphic to a family F_1, where F_1 is obtained as above except that the condition $f(0) < f(1) = 1 < f(2)$ is replaced by the conditions $f(0) = 0$, $f(2) = 2$.

By Theorem I. 9 and the definitions we have the results:

(iii) *Let F be a curve-family filling the set R on K. Let C be a closed curve of F such that F is regular at every point of C. If $C \subset R^i$, then there is a set R_0 such that $F[R_0]$ has C as element and is an open s-family. If $C \subset R - R^i$, then there is a set R_0 such that $F[R_0]$ has C as element and is a half-open s-family.*

1.4. *F* and *R* fixed. From this point on and throughout the paper R will denote the region $K - N - S$, where N and S are two distinct points of K. Without loss of generality, we shall assume N and S to be respectively the north and south poles of K, i. e., $(0,0,1)$ and $(0,0,0)$, and shall use geographical language freely to indicate positions on K: e. g., the equator, latitudinal and longitudinal circles. We shall on occasion use the standard stereographic projection of $K - N$ on π. F will always denote a curve-family filling R and regular in R. K will be assumed oriented in such a way that the positive orientation on the equator is from west to east. π will be assumed oriented so that the counterclockwise direction is positive. The stereographic projection is then a homeomorphism which preserves orientation. Such a homeomorphism will be termed an *o-homeomorphism.*

---

[3] Cf. I, § 1.8. Open *s*-family = *s*-neighborhood.

1. 5. Definition of the sets $L(C,+)$ and $L(C,-)$. By Theorem I. 1 every curve C of F is either open or closed. To an open curve, given parametrically by $p=f(x)$ for $0<x<1$, there can be assigned a positive direction in two ways: with increasing x, or with decreasing x. By a change of parameter, if necessary, we can assume it is with increasing x. A *positive* (*negative*) *limit point* of C shall then mean a point q which is the limit of a sequence $p_n = f(x_n)$ $(n=1,2,\cdots)$, with $\lim x_n = 1$, $(\lim x_n = -1)$. In either case q will be termed a *limit point* of C. We denote by $L(C,+)$ and $L(C,-)$, respectively, the sets of positive and negative limit points of C and by $L(C)$ the set $L(C,+)+L(C,-)$. $L(C,\pm)$ will denote a fixed one of the sets $L(C,+)$, $L(C,-)$. If C is undirected, $L(C)$ is defined uniquely as the set $L(C)$ obtained from either mode of directing C.

From Theorems I. 7, I. 11, I. 9, and I. 13 we have the following theorems:

(iv) *Every set of the form $L(C,+)$, $L(C,-)$ or $L(C)$ is a union of curves of F plus possibly N or S or both.*

(v) *If $L(C,\pm)$ contains a closed curve D, then $L(C,\pm)=D$.*

(vi) *If neither $N \subset L(C,\pm)$ nor $S \subset L(C,\pm)$, then $L(C,\pm)$ contains a closed curve.*

(vii) *Every closed curve D of F separates N and S.*

1. 6. Bays[4] and harbors. Let a curve C of F meet the cross-section pq at successive points t and u interior to pq; let t and u determine arcs $(tu)_1$ on pq and $(tu)_2$ on C meeting only at t and u and hence forming a closed curve Γ. Γ divides the sphere into two regions $\mathfrak{D}(\Gamma)$ and $\mathfrak{D}^*(\Gamma)$ as in Figure 6a or Figure 6b. If p and q are in the same region, as in Figure 6b, Γ is termed a *bay.* If p and q are in different regions, as in Figure 6a, Γ is termed a *harbor.*

---

[4] The definition of bay given here broadens that of I, § 1. 8.

From Theorem I. 12 we have the theorem:

(viii) *If* Γ *is a bay, with* p *and* q *in* $\mathfrak{D}^*(\Gamma)$, *then* $\mathfrak{D}(\Gamma)$ *contains either* N *or* S.

From the lemma of I, § 1. 5 it follows that the cross-section tu lies in an r-neighborhood U and hence that the curves of $F[U]$ can be oriented properly (see I, § 3. 4). We can thus speak of whether C crosses tu at t and u in the same or opposite directions. We then have the result:

(ix) Γ *is a harbor or a bay according as* C *crosses* tu *at* t *and* u *in the same or in opposite directions.*

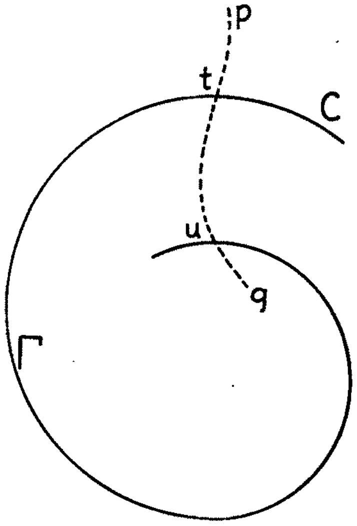


Fig. 6a. Harbor.

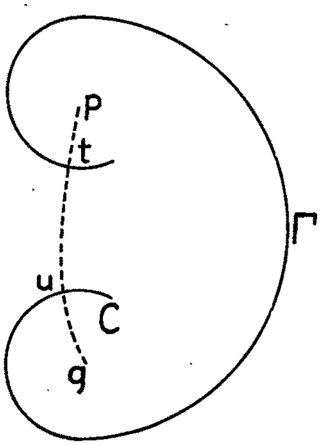


Fig. 6b. Bay.

THEOREM 1. *Let* Γ *be either a bay or a harbor. Then* Γ *separates* N *and* S.

Proof. Let Γ be a bay as in Figure 6b, with $p + q \subset \mathfrak{D}^*(\Gamma)$. By (viii) we can assume that $N \subset \mathfrak{D}(\Gamma)$ and we must then show that $S \subset \mathfrak{D}^*(\Gamma)$. Let w be a point of $(tu)_1$ not on C. The curve D through w then enters $\mathfrak{D}^*(\Gamma)$ in one direction and (a) leaves again at a first point w' or (b) fails to leave $\mathfrak{D}^*(\Gamma)$ again. In case (a) a second bay Γ_0 is formed with $\mathfrak{D}(\Gamma_0) \subset \mathfrak{D}^*(\Gamma)$ and $t + u \subset \mathfrak{D}^*(\Gamma_0)$. Hence by (viii) $S \subset \mathfrak{D}(\Gamma_0) \subset \mathfrak{D}^*(\Gamma)$. In case (b), D is an open curve, and we assume it enters $\mathfrak{D}^*(\Gamma)$ in the positive direction. By (v), $L(D, +)$ then includes either N or S

(hence S) or a closed curve E, where $E \subset \mathfrak{D}^*(\Gamma)$. By (vii) E separates N and S, and hence $S \subset \mathfrak{D}^*(\Gamma)$.

The proof in the case of a harbor is similar.

1. 7. The sets $L(C)$.

Theorem 2. *If C is a directed open curve of F, then $L(C, \pm)$ cannot contain both N and S.*

Proof. Suppose, for example, that $L(C, +)$ contains both N and S. Then by (vii) F can contain no closed curves. Hence by (v) $L(C, -)$ contains either N or S, e. g., S.

C must cross the equator infinitely often in the positive direction, hence it has a limit point p thereon. C must cross a suitably chosen cross-section pq infinitely often in the positive direction. Let t and u be successive crossings, determining a cross-section $(tu)_1$ on pq and an arc $(tu)_2$ on C. Let Γ be the closed curve they form. Γ is then either (a) a bay or (b) a harbor. In either case Γ separates N and S and we choose $\mathfrak{D}(\Gamma)$ to include S.

In case (a) either (a$_1$) $p + q \subset \mathfrak{D}^*(\Gamma)$ or (a$_2$) $p + q \subset \mathfrak{D}(\Gamma)$.

In case (a$_1$) C enters $\mathfrak{D}^*(\Gamma)$ at t and u in the negative and positive directions respectively. But, since $L(C, +) \supset S$ and $L(C, -) \supset S$, C must cross $(tu)_1$ again at first points t' and u' in both directions. Let $(tt')_1 + (tt')_2 = \Gamma_-$. Let $(uu')_1 + (uu')_2 = \Gamma_+$. Then $\Gamma_- \cdot \Gamma_+ = 0$. Choose $\mathfrak{D}(\Gamma_+)$ and $\mathfrak{D}(\Gamma_-)$ so that $\mathfrak{D}(\Gamma_+) \supset S$ and $\mathfrak{D}(\Gamma_-) \supset S$. It follows from Theorem 1 that either $\mathfrak{D}(\Gamma_+) \subset \mathfrak{D}(\Gamma_-)$ or $\mathfrak{D}(\Gamma_-) \subset \mathfrak{D}(\Gamma_+)$. But S can be joined to either t or u without meeting Γ_+ and Γ_- at any other points. This gives a contradiction.

The proof is similar in cases (a$_2$) and (b).

Theorem 3. *If C is a directed open curve of F such that $L(C, +) \supset S$ and $L(C, -) \supset S$, then $L(C) \equiv S$. Similarly $L(C, +) \cdot L(C, -) \supset N$ implies $L(C) \equiv N$.*

The method of proof is the same as for Theorem 2.

Theorem 4. *A directed open curve C of F cannot satisfy the condition $L(C) \equiv D$, where D is a closed curve of F.*

Proof. Suppose $L(C) \equiv D$ and $C \subset \mathfrak{D}(D)$, where $S \subset \mathfrak{D}(D)$. By the Jordan Theorem C divides $\mathfrak{D}(D)$ into two open simply-connected regions $\mathfrak{D}(C)$ and $\mathfrak{D}^*(C)$; we assume $S \subset \mathfrak{D}(C)$. $F_0 = F[\mathfrak{D}^*(C) + C]$ is then regular in $\mathfrak{D}^*(C) + C$. F_0 is thus homeomorphic to a subfamily $F_1[C_1 + \mathfrak{D}(C_1)]$ of a regular curve-family F_1 filling the plane. It follows from

Theorem I. 15 that no curve of F_0 can meet a cross-section twice. But since $L(C) \equiv D$, C must meet infinitely often a cross-section pq with p on D, q in $\mathfrak{D}(D)$. At each intersection of pq with C pq crosses from $\mathfrak{D}(C)$ to $\mathfrak{D}^*(C)$. Hence there must be a cross-section p_1q_1 with p_1q_1 in $C + \mathfrak{D}^*(C)$ and p_1 and q_1 on C. This gives a contradiction.

1. 8. **The three cases.** These results allow us to state that the curves C of F are of eight possible types:

(1) $L(C) = N$ or $L(C) = S$;

(2) $L(C, +) = N$ and $L(C, -) = S$;

(3) $L(C, +) = N$ plus certain open curves of F and $L(C, -) = S$, or $L(C, -) = S$ plus certain open curves of F and $L(C, +) = N$;

(4) $L(C, +) = N$ plus certain open curves of F and $L(C, -) = S$ plus certain open curves of F;

(5) C is closed;

(6) $L(C, +)$ is a closed curve D_1 and $L(C, -)$ is a second closed curve D_2;

(7) $L(C, +)$ is a closed curve D and $L(C, -) = N$ or $L(C, -) = S$;

(8) $L(C, +)$ is a closed curve D and $L(C, -) = N$ or S (not both) plus certain open curves of F.

A curve of type (2) will be said to join N and S. $C + N + S$ is then an arc. If F contains such a curve, then F can obtain no closed curve. F may, however, contain neither a curve of type (2) nor a closed curve. We thus subdivide our problem into three distinct cases:

Case (a). *F contains a curve joining N and S.*

Case (b). *F contains neither a curve joining N and S nor a closed curve.*

Case (c). *F contains a closed curve.*

It will be seen below that in case (a) F consists of curves of types (1) and (2); in case (b) F consists of curves of types (1), (3), and (4); in case (c) F consists of curves of types (1), (5), (6), (7) and (8).

2. Case (a). A curve of *F* joins *N* and *S*.

Throughout this section we assume case (a) holds. Let C denote a curve of F joining N and S.

2.1. Reduction to a family in a strip. If we cut the sphere minus N and S open along C we obtain a set which can be mapped homeomorphically on the strip $|x| \leqq 1$ of the plane, the two images of C being $x = 1$ and $x = -1$. The image of F is a regular family filling the strip and, if we adjoin the lines $x =$ constant outside the strip, we obtain a regular family filling the plane (the regularity being a consequence of (ii) above). Thus F can be obtained from a certain subfamily F_0, which fills a strip, of a regular curve-family F_1 filling the plane by identifying opposite points of the two boundary curves of F_0. Conversely, from any family F_0 which is a subfamily of a regular curve-family filling the plane and which fills a region bounded by two curves C_1 and C_2 of F_0, a family F on the sphere of case (a) can be obtained by identifying opposite points of C_1 and C_2.

The possibilities for the family F_0 are discussed and classified implicitly in I and II, so that we shall consider them no further here.

2.2. Double chordal systems. Let (r, θ, z) be cylindrical coordinates in space. The equations $z' = z$, $\theta' = \theta$, $r' = 1$, define a homeomorphism of $R = K - N - S$ on the interior of the finite cylindrical surface $\mathcal{K}$ whose bounding circles are $\mathcal{N} : r = 1, z = 2$ and $\mathcal{S} : r = 1, z = 0$. This map induces an orientation on $\mathcal{K}$. Under this map F becomes a regular curve-family F^* filling $\mathcal{K}^i$. Each curve of F^* has all its points in $\mathcal{N}$ and $\mathcal{S}$.

Furthermore $\mathcal{K}^i = \mathcal{K} - \mathcal{N} - \mathcal{S}$ can be mapped o-homeomorphically on itself in such a way that F^* becomes a family F^* such that each curve of F^* has a unique limit point $L(C^*, +)$ on $\mathcal{N}$ or $\mathcal{S}$ and a unique limit point $L(C^*, -)$ on $\mathcal{N}$ or $\mathcal{S}$, and no two curves have a common limit point. This map can further be chosen so that, for each curve C^* of F^*, $L(C, \pm) = N$ implies $L(C^*, \pm) \subset \mathcal{N}$ and $L(C, \pm) = S$ implies $L(C^*, \pm) \subset \mathcal{S}$.

The existence of this map can be established as follows: Proceed as in § 2.1 to obtain the families F_0 and F_1. By a homeomorphism reduce F_1 to a family filling the interior of a circle and F_0 to a sub-family bounded by two parallel chords C_1 and C_2. By Theorem II. 29 this homeomorphism can be so chosen that each curve of F_1 has a unique limit point in each direction on the circumference and no two curves have a common limit point. By identifying corresponding points of C_1 and C_2 and a subsequent homeomorphism we then obtain the family F^* on the cylinder $\mathcal{K}$ as desired.

Now replace each curve of F^* by the line segment joining its two limit points. We then obtain a system of line segments E such that each line of E joins a point of $\mathcal{N}$ or $\mathcal{S}$ to a point of $\mathcal{N}$ or $\mathcal{S}$ and no two segments have an end-point in common. Such a system of lines or "chords" will be termed a *double chordal system* (cf. the chordal systems of I and II). The particular double chordal system arising from F will be denoted by $DCS(F)$. This is not uniquely determined.

Two double chordal systems E_1, E_2 will be termed *isomorphic* if there is a one-to-one correspondence between them such that each chord of $\mathcal{S}$ corresponds to one of $\mathcal{S}$, each chord of $\mathcal{N}$ to one of $\mathcal{N}$, each chord of $\mathcal{S}$ and $\mathcal{N}$ to one of $\mathcal{S}$ and $\mathcal{N}$, and such that any three chords of the first system have the same relative positions as the corresponding three of the second system. Any two double chordal systems arising from the same family F must then be isomorphic; hence $DCS(F)$ is unique up to an isomorphism.

However, whereas there are only five possible relative positions of three chords in the case of the ordinary chordal systems (cf. I § 2 Fig. 2) here there are sixty possibly relative positions. Two different arrangements, in the case of the double chordal systems, have to be distinguished by the relative positions of their end-points, as two chords may intersect.

2. 3. Classification in case (a). *Definition.* Let F be a regular curve family filling R. By the *four forms* of F will be meant the four curve-families F_i $(i = 1, 2, 3, 4)$ filling R which are the respective images of F under the following transformations: $T_1 =$ the identity; $T_2 =$ a reflection in the equator; $T_3 =$ a reflection in the great circle of K in the xz-plane; $T_4 = T_3T_2 = T_2T_3$.

The concept of a normal explicit chordal system was defined in I § 3. 3 and II § 3. 2. If E is a normal explicit chordal system and C_1 and C_2 are two parallel chords of E, then by identifying opposite points of C_1 and C_2 and mapping the zone of C_1 and C_2 homeomorphically on $\mathcal{K}$, as in § 2. 2, the subset $\lambda(C_1 + C_2)$ of E becomes a double chordal system on $\mathcal{K}$. A double chordal system obtained in this way from a normal chordal system will be termed *normal.*

The preceding analysis and the results of I and II then permit us to state the following theorem:

THEOREM 5. *Let F_1 and F_2 be two regular curve-families of case (a) filling R. Then F_1 is homeomorphic to F_2 if and only if for suitably chosen forms F'_1 and F'_2 of F_1 and F_2 every $DCS(F'_1)$ is isomorphic to every $DCS(F'_2)$. Furthermore a double chordal system E is of the form $DCS(F)$ for some regular curve-family F filling R if and only if E is normal.*

3. Case (b). *F* contains no closed curve or joining curve.

We assume throughout this section that case (b) holds: F contains neither a closed curve nor a curve joining N and S.

Throughout the following discussion a number of statements will be made which apply equally to N or to S. This will be indicated by the expressions "similarly for S" or "similarly for N," implying that S and N are to be interchanged in the statement.

3.1. **Flowers.** Let C be a curve of F such that $L(C) = N$. $C + N$ is then a closed curve and divides K into two regions. We shall always denote by $\mathfrak{D}(C)$ that region which does not contain S and by $\mathfrak{D}^*(C)$ the other region. In both cases only interior points are included. (Similarly for S).

Let C_1 and C_2 be two curves of F such that $L(C_i) = N$ $(i = 1, 2)$. We write $C_1 \sim C_2$ if $\mathfrak{D}(C_1) \cdot \mathfrak{D}(C_2) \neq 0$, $C_1 \not\sim C_2$ if $\mathfrak{D}(C_1) \cdot \mathfrak{D}(C_2) = 0$. (Similarly for S).

Definition. A *flower* at N is a subfamily Fl of F such that for each curve C in Fl $L(C) = N$ and, if C_1 and C_2 are in Fl, then $C_1 \not\sim C_2$. (Similarly for S). (See Figure 2 in the Introduction).

THEOREM 6. *A flower at N or S is an at most countable collection of curves.*

THEOREM 7. *If $Fl \equiv \{C_n\}$ $(n = 1, 2, \cdots)$ is an infinite flower at N, then the curves C_n approach N uniformly as $n \rightarrow \infty$. (Similarly for S).*

Remark. The conclusion of the theorem is understood to mean the following: if p_n is any sequence of points, with $p_n \subset C_n$, then $\lim_{n \to \infty} p_n = N$.

Proof of theorem. Suppose the theorem false. Then there must be a point p which is a limit point of a sequence $\{p_n\}$, with p_n on C_n, $p \neq N$, $p \neq S$. Draw a cross-section ab through p. ab must cross infinitely many C_n. At each crossing of a C_n ab must go from $\mathfrak{D}^*(C_n)$ to $\mathfrak{D}(C_n)$. Since the $\mathfrak{D}(C_n)$ are disjoint, this implies that a bay Γ will be formed with $\mathfrak{D}(\Gamma) \subset \mathfrak{D}(C_n)$ for some n. This contradicts Theorem 1.

THEOREM 8. *If C is a directed open curve of F for which $L(C, +) \supset N$ but $L(C, +) \not\equiv N$, then $L(C, +)$ is a flower at N. (Similarly for S).*

Proof. $L(C, +)$ consists of N plus certain curves $\{D_a\}$, by (iv). $L(C, +)$ is a closed set and (by Theorem 2) does not contain S. Hence no $L(D_a, \pm)$ can contain S. Since F has no closed curves, it follows from the classification at the end of § 1 that for each D_a $L(D_a) = N$.

If D_1 and D_2 are two curves of $L(C, +)$, then $D_1 \sim D_2$ would imply $C \subset \mathfrak{D}(D_1)$ or $C \subset \mathfrak{D}(D_2)$. But by Theorem 3 $L(C, -) \supset S$, which is not possible if $C \subset \mathfrak{D}(D_1)$ or $C \subset \mathfrak{D}(D_2)$. Hence $D_1 \not\sim D_2$, and thus the curves $\{D_a\}$ form a flower at N.

THEOREM 9. *Let* $Fl = \Sigma C_n$ *be a flower at* N. *Then the set* $M = \prod \mathfrak{D}^*(C_n)$ *is an open simply-connected region whose boundary is* $N + \Sigma C_n$. *(Similarly for* S*).*

The theorem is proved by considering the possible boundary of a component of M.

3.2. The sets $B(N)$ and $B(N, C)$. *Definition.* $B(N)$ will denote the family of all curves C of F such that $L(C) \equiv N$.

We partially order $B(N)$ thus: $C_1 < C_2$ if $\mathfrak{D}(C_1) \subset \mathfrak{D}(C_2)$. If C is in $B(N)$ we let $B(N, C)$ denote the subfamily of $B(N)$ consisting of all elements of $B(N)$ greater than C (equality included). $B(N, C)$ is then simply ordered. (Similarly for S).

THEOREM 10. *For every curve* C_0 *in* $B(N)$, $B(N, C_0)$ *has a greatest element. (Similarly for* S*).*

Proof. Suppose the theorem to be false. Then the set $G = \Sigma \mathfrak{D}(C)$ for C in $B(N, C_0)$ is an open simply-connected set and the point set $B(N, C_0) \subset G$. The boundary of G consists of N, possibly S, and certainly other points. Let Q be a boundary point of G, with $N \neq Q \neq S$. Then, because of the regularity of F (see I § 1.7) the whole curve D of F through Q is on the boundary of G. The boundary thus consists of curves of F plus N and possibly S.

Let D be any boundary curve of G. Since the boundary is closed, $L(D)$ is in the boundary. But we have

$$L(D) \subset N + S. \tag{1}$$

For suppose $L(D)$ contains a curve D_1. Then choose an r-neighborhood $U(Q_1)$ of a point Q_1 on D_1. $U(Q_1)$ is then divided into two smaller r-neighborhoods U' and U'' by an arc δ on D_1. Since D_1 is in the boundary of G and G is a union of curves of $B(N)$, it follows (with the help of Theorem 1) that, if δ is a limit of arcs in U' of curves C of G, then $U' \subset G$, and similarly for U''. This contradicts the fact that D_1 is in $L(D)$ for the boundary curve D. Thus equation (1) follows.

Since we assume case (b), we must have $L(D) = N$ or $L(D) = S$; i.e., $D \subset B(N)$ or $D \subset B(S)$. Since G is connected, the boundary curves D in $B(N)$ form a flower at N, and similarly at S. By Theorem 7 each of these flowers plus N or S is a closed set. Since G is simply-connected, only one of the flowers can be in the boundary. Hence the boundary consists of a flower Fl in $B(N)$ plus N itself. Let $Fl = \Sigma D_n$. Then, by Theorem 9, $G = \prod \mathfrak{D}^*(D)$, and hence $G \supset S$. This is impossible. Hence $B(N, C_0)$ has a largest element and Theorem 10 is established.

3. 3. Decomposition of F into $B(N)$, $B(S)$, and $J(F)$. For each element C of $B(N)$ we now let C^* denote the largest element of $B(N, C)$. Let G^* be the set of all such C^*. G^* must then be a flower and we denote it by $Fl(N)$. Every curve C of $B(N)$ is then less than or equal to some curve of $Fl(N)$. We write $\mathfrak{D}Fl(N)$ for the set $\Sigma\mathfrak{D}(C)$, C in $Fl(N)$. (Similarly for S).

Now let $J(F) = F - B(N) - B(S)$. Each curve of J then satisfies $L(C, +) \subset Fl(N) + N$, $L(C, -) \subset Fl(S) + S$ for proper orientation of C.

We have thus decomposed the family F into three pieces $B(N)$, $B(S)$, and $J(F)$. Each of $B(N)$, $B(S)$ is a flower plus interior. $J(F)$ consists of curves 'joining' $B(N)$ and $B(S)$. In the following sections we shall analyze the internal structure of each of the pieces and determine how they fit together. It will turn out that $B(N)$ and $B(S)$ have the structure of a subset $E - \theta(V)$ of a normal chordal system E, while $J(F)$ has the structure of the family of longitudinal semi-circles. The curves of $J(F)$ can all wind around $B(N)$ like spirals or they can all tend directly to N, and similarly at S. In the latter case the picture of the singularity at N, or S, is precisely that of a regular family filling the plane at the point at infinity.

3. 4. Structure of $J(F)$.

THEOREM 11. *The curves of $J(F)$ fill an open doubly-connected region Φ. Φ can be mapped homeomorphically on R in such a way that the family $J(F)$ becomes the family of (open) longitudinal semi-circles.*

Proof. That $\Phi = R - (B(N) + B(S))$ is doubly-connected follows by repeated application of Theorem 9. Since every curve C of $J(F)$, when properly oriented, satisfies $L(C, +) \supset N$, $L(C, -) \supset S$, Φ can be mapped homeomorphically on R so that each curve C' of the image family J' satisfies $L(C', +) = N$, $L(C', -) = S$. J' is then a regular curve-family filling R and one to which case (a) applies. It follows immediately from Theorem 5 that J', and hence $J(F)$, is homeomorphic to the family of longitudinal semicircles in R.

COROLLARY. *There exists a closed cross-section in Φ, i. e., a closed curve γ lying in Φ which meets each curve of $J(F)$ exactly once; further, γ can be chosen to separate N and S.*

For let T_0 be the map of $J(F)$ on the longitudinal semi-circles; let g be any latitudinal circle. $T_0^{-1}(g)$ satisfies the conditions of the corollary.

Let s be the first point at which C' meets pw. The arcs $(ws)_1$ on pq and $(ws)_2$ on C' then form a harbor Γ. Let Z denote the closed region of the plane bounded by Γ and $r = 1$. Cut Z open along psw and then map it homeomorphically on the rectangle I: $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$ so that $r = 1$ becomes $y = 0$, $(ws)_2$ becomes $y = 1$, p has inverses $(0, 0)$ and $(1, 0)$, s has inverses $(0, y_0)$ $(0 < y_0 < 1)$ and $(1, 1)$. The family in the annulus becomes a family ψ in I consisting of arcs joining a point of $x = 0$, $0 < y \leqq 1$, to a point of $x = 1$, $0 < y \leqq 1$ plus open intervals C''_n in $0 < x < 1$, $y = 0$ $(n = 2, 3, \cdots)$ and two half-open intervals C''_1. We denote by η the sets of points of $y = 0$, $0 \leqq x \leqq 1$, not on a C''_n.

LEMMA A. *There is a homeomorphism T_1 of the set $I - \eta$ on itself such that $T_1(C''_n) = C''_n$ for each n, each point of I on $y = 1$ is fixed, while the curves of ψ in $y > 0$ become lines $y = constant$.*

Remark. This lemma may be interpreted as a regularization of ψ at the points of η.

The lemma is established by imbedding ψ minus the points of $x = 0$ and $x = 1$ in a regular curve-family filling the interior of a circle, then doing the same with the family obtained from ψ by replacing the curves in $y > 0$ by lines $y =$ constant and then applying Theorem II. 27.

We now apply T_1 and then map I back on Z by a homeomorphism which identifies points on $x = 0$ and $x = 1$ corresponding ultimately to the same point of the annulus. Thus the family F' has now been regularized along $r = 1$; i. e., F' remains regular if the C'_n are replaced by the circle $r = 1$. Let F'_+ denote the regular family thereby obtained.

We next demonstrate that F'_+ is homeomorphic with the family χ consisting of the circle $r = 1$ plus the spiral curves

$$\theta = \tan \frac{\pi(2r - 3)}{2} + k, \qquad 1 < r \leqq 3/2,\ 0 \leqq k \leqq 2\pi.$$

To that end we first extend the cross-section pq to become a larger one pu, where u lies on $r = 3/2$ and $pu - u$ lies in $\mathfrak{D}(\gamma)$. That this is possible follows from Theorem 11.

Let C' be the curve of F'_+ through u. C' then meets pu at a first point v after u. Let Γ denote the harbor thereby formed. Let Z denote the closed region bounded by Γ and $r = 1$ and let Z^* denote the remainder of the annulus.

If we cut open along pv then Z can be mapped exactly as was the Z above on the rectangle I so that the curves of $F'_+[Z]$ become the lines $y =$ constant. There is a homeomorphism T_0 corresponding, which identifies points on $x = 1$ and points on $x = 0$. T_0 can be regarded as a function of y with values y, and we have $T_0(0) = 0$, $T_0(1) = y_0 < 1$. Further $T_0(y) \neq y$ except for $y = 0$, and hence in general $T_0(y) < y$. By a suitable further homeomorphism of the rectangle on itself leaving the family unchanged T_0 can be made to coincide with any homeomorphism with the above properties. This follows from a lemma on monotone functions.[5]

We now proceed in exactly the same way with the family χ. We choose the cross-section $p_\chi v_\chi$ as the line joining $(1, 0^0)$ and $(3/2, 0^0)$. We choose the curve Γ_χ through v_χ and u_χ in the same way as above and obtain the regions Z_χ and Z^*_χ. We cut open along $p_\chi v_\chi$ and map Z_χ on the rectangle I so that the curves of $\chi[Z_\chi]$ become parallel lines $y = \text{constant}$. We have an identification $T_{0\chi}$ of $x = 1$ and $x = 0$. But we can assume, by the preceding paragraph, that $T_{0\chi} \equiv T_0$. It thus follows that there is a homeomorphism T of $F'_+[Z]$ on $\chi[Z_\chi]$ taking Γ on Γ_χ, $r = 1$ on $r = 1$. T can then be extended to a homeomorphism of all F'_+ on all χ by cutting Z^* and Z^*_χ open at the points u and u_χ to get triangles and then proceeding as above. We thus conclude that there is a homeomorphism of $F[W]$ on the family χ from which points on $r = 1$ have been removed. Under this map the curves C_n of $Fl(N)$ become arcs C'_n on $r = 1$. To complete the proof of Theorem 13 it remains only to extend the homeomorphism to each set $\mathfrak{D}(C_n)$ in such a way that each is mapped on the interior of the segment of the circle bounded by C'_n. Thus Theorem 13 is established.

3.7. Relationship between $J(F)$, $B(N)$ and $B(S)$. Case 2.

Throughout § 3.7 we assume that case 2 holds: $J(F)$ contains a curve C of F such that $L(C, +) = N$.

We now define a family of curves on K to be *of type* (ϵ) if it is homeomorphic to a family $F^0[\mathfrak{D}(\gamma^0) - N]$, where F^0, γ^0, and $\mathfrak{D}(\gamma^0)$ satisfy the following conditions:

(ϵ_1) F^0 is a regular curve-family filling R of case (a);

(ϵ_2) γ^0 is a closed cross-section relative to F^0 (cf. corollary to Theorem 11);

(ϵ_3) γ^0 divides K into two regions, one of which, $\mathfrak{D}(\gamma^0)$, contains N, while the other, $\mathfrak{D}^*(\gamma^0)$, contains S;

---

[5] LEMMA. *Let $g_i(y)$ be continuous monotone strictly increasing functions of y for $a \leqq y \leqq b$, with $g_i(a) = a$ and $g_i(y) < y$ for $a < y \leqq b$ $(i = 1, 2)$. Then there is a homeomorphism T of the interval $[a, b]$ on itself such that $g_2(y) = Tg_1(T^{-1}(y))$.*

The lemma is proved by subdividing $[a, b]$, first for $i = 1$ and then for $i = 2$, by the points $g_i(b), g_i(g_i(b)), \cdots$ and then determining a suitable map of each interval of one subdivision on the corresponding interval of the second subdivision.

(ϵ_4) for every point p in $\mathcal{D}^*(\gamma^0)$, the curve C_p of F^0 through p joins N and S.

THEOREM 14. *In case 2 for F the family $F[\gamma + \mathcal{D}(\gamma) - N]$ is of type* (ϵ).

Proof. From the regularity of F and the fact that every arc on γ is a cross-section it follows that $F[\gamma + \mathcal{D}(\gamma) - N]$ is regular at each point of γ. Now map $\gamma + \mathcal{D}(\gamma)$ homeomorphically on the northern hemisphere so that γ becomes the equator γ^0 and N remains fixed. Adjoin the longitudinal quarter-circles joining the points of the equator to S. It follows from (ii) (§ 1. 2) that the resulting family F^0 in R is regular at every point of R. Since F contains a curve C of $J(F)$ with $L(C, +) = N$, F^0 must contain a curve C^0 joining N and S; in fact (ϵ_4) must then hold. Thus F^0 is of case (a). If we take $\mathcal{D}(\gamma^0)$ to be the northern hemisphere minus γ^0, then $F[\gamma + \mathcal{D}(\gamma) - N]$ is homeomorphic to $F^0[\gamma^0 + \mathcal{D}(\gamma^0) - N]$. Since ($\epsilon_1$), ($\epsilon_2$), ($\epsilon_3$), ($\epsilon_4$) all hold, the theorem is established.

Remarks. It follows from Theorem 5 that, given any family of type (ϵ), the corresponding family F^0 is unique up to a homeomorphism. Hence the families of type (ϵ) are automatically classified by Theorem 5; they are in one-to-one correspondence with a subclass of all regular families of case (a) filling R, whereby homeomorphic families are considered as the same family. This subclass can be characterized in terms of double chordal systems (see § 2. 2): a family belongs to the subclass if and only if it is represented by a double chordal system E on $\mathcal{N}$ and $\mathcal{S}$ such that the subsystem of E consisting of those chords of E which join $\mathcal{N}$ and $\mathcal{S}$ has the same structure as the family of rulings of $\mathcal{K}$, while no chord of E joins two points of $\mathcal{S}$.

3. 8. Classification for case (b). A full classification requires three stages: (1) a listing of all cases which arise; (2) a proof that the cases of (1) actually do arise; (3) a setting up of criteria for the topological equivalence of two families.

For stage (1) Theorems 11, 12, 13 and 14 give necessary conditions for the families. These state that the structure of $J(F)$ is uniquely determined up to a homeomorphism, that the structure of $B(N)$ is that of a cyclic complement, that F can be decomposed into two pieces $F[\gamma + \mathcal{D}(\gamma) - N]$ and $F[\gamma + \mathcal{D}^*(\gamma) - S]$, in each of which two structures, (δ) and (ϵ), are possible. The (δ) structure is that of spirals about a closed curve; the (ϵ) structure is essentially a special case of that appearing in case (a). Considering possible

n14905

combinations in $\mathfrak{D}(\gamma)$ and $\mathfrak{D}^*(\gamma)$, we can group the families of case (b) in two types:

Type $(N\delta, S\delta)$: $B(N) \neq 0$, $B(S) \neq 0$ and $F[\gamma + \mathfrak{D}(\gamma) - N]$ and $F[\gamma + \mathfrak{D}^*(\gamma) - S]$ are both of type (δ).

Type $(N\delta, S\epsilon)$: $B(N) \neq 0$ and $F[\gamma + \mathfrak{D}(\gamma) - N]$ is of type (δ); $F[\gamma + \mathfrak{D}^*(\gamma) - S]$ is of type (ϵ).

We note that since, in case (b), no curve of F can join N and S, there is no third type, $(N\epsilon, S\epsilon)$.

For the type $(N\delta, S\delta)$ there is an immediate further subdivision to be made. For the homeomorphism of $F[\gamma + \mathfrak{D}(\gamma) - N]$, for example, on the family of spirals may or may not be an o-homeomorphism. In case it is an o-homeomorphism, we shall say that the family $F[\gamma + \mathfrak{D}(\gamma) - N]$ is of type $(\delta +)$, otherwise that it is of type $(\delta -)$. The possible combinations at N and S are then $(N\delta +, S\delta +)$, $(N\delta -, S\delta -)$, $(N\delta +, S\delta -)$, $(N\delta -, S\delta +)$. We shall refer to these as the *four subtypes.* We note that an orientation reversing homeomorphism of K on itself leaving N and S fixed interchanges the first two combinations and also the last two.

Similarly, for the type $(N\delta, S\epsilon)$ we can distinguish between the two subtypes $(N\delta +, S\epsilon)$ and $(N\delta -, S\epsilon)$.

We now define a family A of curves on the sphere K to be an *admissible* $B(N)$ if A is homeomorphic to a cyclic complement F'_0 under a stereographic projection of $K - N$ on the xy-plane. (Similarly for S).

THEOREM 15. *Let A_1 be an admissible $B(N)$, let A_2 be an admissible $B(S)$, and let $A_1 \cdot A_2 = 0$. Then there exists a regular curve-family F filling R, of case (b), for which $B(N) = A_1$, $B(S) = A_2$, and F is of type $(N\delta, S\delta)$. F can be chosen to be of any one of the subtypes of type $(N\delta, S\delta)$.*

Proof. Map the open region bounded by N, S, and $Fl(N)$ and $Fl(S)$ on the annulus $1 < r < 2$ by a homeomorphism T which can be extended to the boundary curves of $Fl(N)$ and $Fl(S)$. $T(Fl(N))$ is then a set of open arcs, e.g., on $r = 1$; $T(Fl(S))$ is a set of open arcs on $r = 2$. In the annulus $1 < r < 2$ place one of the four families:[6] (1) the family of Figure 5a (see Introduction, § 6); (2) the family obtained from (1) by reflecting

---

[6] These can be defined analytically as follows:

$$\theta = \epsilon \tan [\pi(r - 3/2)] + k$$
$$\theta = \epsilon \sec [\pi(r - 3/2)] + k$$

for $1 < r < 2$ and $0 \leq k \leq 2\pi$. $\epsilon = +1$ gives (1) and (3); $\epsilon = -1$ gives (2) and (4).

in a diameter of $r = 2$; (3) the family of Figure 5b; (4) the family obtained from (3) by reflecting in a diameter of $r = 2$. The inverse images of these families then form with A_1 and A_2 regular families F filling R of case (b) and they give the four subtypes as desired.

It thus follows that the classification requirements (1) and (2) for type $(N\delta, S\delta)$ have been satisfied. For type $(N\delta, S\epsilon)$ we proceed similarly:

Definition. A family A of curves in R will be termed an *admissible* $S\epsilon$ family if A is of type $(\epsilon)'$ and has S as isolated singularity.

THEOREM 16. *Let A_1 be an admissible $S\epsilon$ family and let A_2 be an admissible $B(N)$ such that $A_1 \cdot A_2 = 0$. Then there exists a regular curve-family F filling R, of case (b), type $(N\delta, S\epsilon)$, for which $B(N) = A_2$ and for which there is a closed cross-section γ in $J(F)$ such that $A_1 = F[\gamma + \mathfrak{D}(\gamma) - S]$, where $\mathfrak{D}(\gamma) \supset S$. F can be chosen to be of either subtype.*

Proof. A_1 will be bounded by a closed cross-section γ and will fill $\mathfrak{D}(\gamma) - S$. The region between A_1 and A_2 can thus be mapped on an annulus and hence essentially the same device as was used in the proof of Theorem 15 can be used to establish the theorem.

It now remains to carry out stage (3) of the classification: the setting up of criteria for the homeomorphism of two families. We use again here the notion of the four forms of a family (see § 2. 3).

Consider now two families F_1 and F_2 of case (b). A necessary condition that they be homeomorphic is that both be of the same type. We can therefore discuss each type separately.

Consider type $(N\delta, S\delta)$. Here we consider first the families $B_i(N)$ and $B_i(S)$ $(i = 1, 2)$. These are all homeomorphic to cyclic complements, the $B_i(N)$ under stereographic projection, the $B_i(S)$ under reflection in the equator followed by a stereographic projection. The *isomorphism* of two of these families is then defined to mean that of the coresponding planar families (regarded as chordal systems, as in II). If $B_1(N)$ and $B_2(N)$ are isomorphic, then $Fl_1(N)$ and $Fl_2(N)$ will necessarily correspond to each other under the isomorphism (similarly for S).

THEOREM 17. *Let F_1 and F_2 be two curve-families of case (b), type $(N\delta, S\delta)$. Then a necessary and sufficient condition that F_1 and F_2 be homeomorphic is that for one form F'_1 of F_1 and one form F'_2 of F_2 the following holds: $B'_1(N)$ is isomorphic to $B'_2(N)$; $B'_1(S)$ is isomorphic to $B'_2(S)$; F'_1 and F'_2 are of the same one of the four subtypes.*

Proof. The necessity is immediate.

For the sufficiency we choose closed cross-sections γ_1 and γ_2 in $J(F'_1)$ and $J(F'_2)$ respectively and choose $\mathfrak{D}(\gamma_1)$ and $\mathfrak{D}(\gamma_2)$ to include N. Both families $F'_1[\gamma_1 + \mathfrak{D}(\gamma_1) - N]$ and $F'_2[\gamma_2 + \mathfrak{D}(\gamma_2) - N]$ are of type (δ) and can be mapped on families F_0 of the definition (see § 3. 6). Since F'_1 and F'_2 are of the same subtype, either both of these homeomorphisms will preserve orientation or both will reverse it. Now $B'_1(N)$ and $B'_2(N)$ are isomorphic under a one-to-one correspondence which necessarily takes $Fl'_1(N) = \Sigma C'_{1n}$ on $Fl'_2(N) = \Sigma C'_{2n}$. It follows that there is a one-to-one correspondence between their images ΣC^*_{1n} and ΣC^*_{2n} which preserves order and orientation on the circle. It thus follows that the ΣC^*_{1n} can be assumed identical with the ΣC^*_{2n}. For we can reduce a neighborhood of $r = 1$ to a rectangle and proceed as in the proof of Lemma A of Theorem 13. This gives an o-homeomorphism of

$$F'_1[\gamma_1 + \mathfrak{D}(\gamma_1) - \Sigma\mathfrak{D}(C'_{1n})]$$

on

$$F'_2[\gamma_2 + \mathfrak{D}(\gamma_2) - \Sigma\mathfrak{D}(C'_{2n})]$$

taking γ_1 on γ_2 and each C'_{1n} on the corresponding C'_{2n}. This o-homeomorphism can be extended by similar methods to give an o-homeomorphism of F'_1 on F'_2. Thus F_1 and F_2 are homeomorphic.

THEOREM 18. *Let F_1 and F_2 be two curve-families of case (b), type $(N\delta, S\epsilon)$. Then a necessary and sufficient condition that F_1 and F_2 be homeomorphic is that for one form F'_1 of F_1 and one form F'_2 of F_2 the following holds: the families $B'_1(N)$ and $B'_2(N)$ are isomorphic; there are closed cross-sections γ_1, γ_2 in $J(F'_1)$ and $J(F'_2)$ such that, if $\mathfrak{D}(\gamma_i) \supset S$ $(i = 1, 2)$, then there is an isomorphism between $DCS(F'_1[\mathfrak{D}(\gamma_1) - S])$ and $DCS(F'_2[\mathfrak{D}(\gamma_2) - S])$; F'_1 and F'_2 are of the same subtype.*

Remark. The $DCS(F'_i[\mathfrak{D}(\gamma_i) - S])$ is understood to mean $DCS(\phi_i)$, where ϕ_i is a curve family filling R obtained from $F'_i[\mathfrak{D}(\gamma_i) - S]$ by an o-homeomorphism of $\mathfrak{D}(\gamma_i)$ on $K - N$ leaving S fixed. $DCS(\phi_i)$ is defined in § 2. 2.

The proof of the theorem is similar to that of Theorem 17.

4. Case (c). *F* contains closed curves.

Throughout this section we assume that case (c) holds: F contains at least one closed curve.

4. 1. Annular families. By (vii) (see § 1. 5) each closed curve of F separates N and S. It follows that of each pair D_1, D_2 of closed curves one separates the other from N. The closed curves can thus be simply ordered by the relationship of being nearer to N, and hence lie on the sphere essentially as a set of latitudinal circles. (It will follow from the results below that a suitable homeomorphism of K on itself transforms them onto latitudinal circles). It will accordingly be necessary to analyze the structure of F in the region between two closed curves. Such a region is homeomorphic to an annulus.

Definition. *A closed annular family* shall mean a curve-family ϕ on the sphere or in the plane which has the following properties: ϕ fills a closed region A whose boundary consists of two non-intersecting closed curves D_1 and D_2; D_1 and D_2 are elements of ϕ; ϕ is regular at every point of A. An open (half-open) annular family shall mean a family which can be written as a sum $\sum_{i=-\infty}^{\infty} \phi_i$ $(\sum_{i=1}^{\infty} \phi_i)$ of closed annular families ϕ_i, whereby $\phi_i \cdot \phi_{i+1}$ is a common bounding curve of ϕ_i and ϕ_{i+1} for each i and $\phi_i \cdot \phi_j = 0$ for $| i - j | > 1$.

The word 'closed' will generally be omitted in referring to a closed annular family. Throughout § 4. 1 there will be no reference to the open or half-open families.

From this point on ϕ will denote a fixed annular family. ϕ may be on the sphere or in the plane but, if it is on the sphere, it will be assumed not to contain N or S.

Definition. ϕ will be termed *simple* if ϕ contains no closed curves besides D_1 and D_2.

If ϕ is simple, then $\phi - D_1 - D_2$ can be mapped homeomorphically on $K - N - S$, and all the reasoning of case (b) can be applied. We conclude, as in the proof of Theorem 16, that ϕ can be mapped o-homeomorphically on one of the four annular families:[7]

(*) the family of Figure 5b;

(**) the family obtained from (*) by reflecting in a diameter of the larger circle;

(***) the family of Figure 5a;

(****) the family obtain from (***) by reflecting in a diameter of the larger circle.

If ϕ is on the sphere, the o-homeomorphism can be accomplished by first stereographically projecting on the plane by a projection P and then performing a further o-homeomorphism. This second o-homeomorphism can be so chosen that the outer boundary curve of $P\phi$ becomes the outer circle of the final image. It will then follow that ϕ can be mapped in this way on exactly one of the four families (*), (**), (***), (****). We term that one the *associated canonical family* of ϕ. If ϕ is already in the plane, then replace P by the identity transformation to obtain the definition of the associated canonical family.

---

[7] See footnote 6 above for the equations of the spirals.

Definition. If the associated canonical family of the simple annular family ϕ is (***) or (****), ϕ will be termed *ordinary*. If the associated canonical family is (*) or (**), ϕ will be termed *bent.*

Now suppose that ϕ contains closed curves besides D_1 and D_2. Since $A - D_1 - D_2$ is homeomorphic to a family filling R, we conclude that the closed curves of ϕ lie as concentric circles between D_1 and D_2; i. e., they can be simply ordered with respect to nearness to D_1 (or to D_2). ϕ may then contain sub-families which are simple annular families. Two such sub-families, if non-coincident, can have at most a bounding closed curve in common.

THEOREM 19. *ϕ contains at most a finite number of simple annular sub-families of bent type.*

Proof. If ϕ contained infinitely many bent sub-families they would have a limiting closed curve C. By (iii) (§ 1. 2) C lies in an s-family. An s-family can contain no bent annular families, and hence there is a contradiction.

Definition. An annular family is *smooth* if it contains no bent sub-families.

COROLLARY TO THEOREM 19. *Every annular family ϕ can be decomposed in a unique way into a sum of annular families as follows:* $\phi = \phi_1 + \phi_2 + \cdots + \phi_n$; $\phi_i \cdot \phi_{i+1}$ *is a closed curve* C_i *of* ϕ $(i = 1, 2, \cdots, n-1)$; $\phi_i \cdot \phi_j = 0$ $(j > i+1)$; *each* ϕ_i *is either bent or smooth; if* ϕ_i *is smooth, then* ϕ_{i+1} *is bent.*

ϕ is thus composed of successive rings of smooth or bent type, the smooth ones being separated by bent ones. To complete the analysis, we must determine the possibilities for a smooth annular family.

THEOREM 20. *If ϕ is a smooth annular family, then ϕ is a closed s-family.*

Proof. From a point p on the boundary curve D_1 of ϕ it is possible to draw a cross-section pq, with q on a closed curve interior to ϕ or on D_2. If it were impossible to have q on D_2, there would certainly be a 'least upper bound' of the closed curves on which q can be chosen: i. e., a closed curve C

such that q can be chosen on any closed curve between D_1 and C. But a cross-section pq to a point q in the neighborhood of C can be extended to become one which crosses C, as follows from (iii) (§ 1. 2). Hence there is a contradiction and q can be chosen on D_2. If we now cut open on pq, we obtain a rectangular family with identifications as in § 1. 3. Thus ϕ is a closed s-family.

COROLLARY 1. *If ϕ is a smooth annular family, then there is a cross-section pq with p on D_1 and q on D_2.*

COROLLARY 2. *If ϕ is a smooth annular family, then ϕ is orientable.*

(See I § 3. 4 for the definition of orientability of a curve-family).

It thus remains to discuss the possibilities for a family ϕ which is a closed s-family. But such a family ϕ is obtained from a rectangle $M : 0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$ and the 'curves' $y = \text{constant}$ filling it by an identification $y' = T(y)$ between the points $(1, y)$ of $x = 1$ and those $(0, y')$ of $x = 0$. $T(y)$ is here a homeomorphism and $T(0) = 0$, $T(1) = 1$. T can thus also be regarded as an o-homeomorphism of the interval $0 \leqq y \leqq 1$ on itself. T is not uniquely determined by ϕ.

Definition. Let T_1 and T_2 be two o-homeomorphisms of $0 \leqq y \leqq 1$ on itself. T_1 and T_2 will be termed *similar* if the following two conditions hold:

(1) If E_i is the (necessarily closed) set of fixed points of T_i $(i = 1, 2)$, then there is a homeomorphism $y_2 = f(y_1)$ of E_1 on E_2, where f is a monotone increasing function.

(2) If y'_i and y''_i belong to E_i, but $y \not\subset E_i$ for $y'_i < y < y''_i$ and $f(y'_1) = y'_2$, $f(y''_1) = y''_2$, then $T_1(y) - y$ and $T_2(y) - y$ have the same sign in the respective intervals $y'_i < y < y''_i$ $(i = 1, 2)$.

The relation of similarity is reflexive, symmetric and transitive, hence is a valid equivalence relation.

The following lemma can then be established:[8]

LEMMA. *Let G_1 and G_2 be two curve-families obtained from the family $y = \text{constant}$ in $M : 0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$, by respective identifications $y' = T_1(y)$ and $y' = T_2(y)$ of the line $x = 1$ on the line $x = 0$ $(0 \leqq y \leqq 1)$, whereby T_i is an o-homeomorphism of the interval $0 \leqq y \leqq 1$ on itself. Then there exists an o-homeomorphism T_0 of G_1 on G_2, taking the line $y = 1$ on the*

---

[8] The proof is similar to that for the lemma cited in footnote 5 above.

line $y = 1$, $y = 0$ *on* $y = 0$ $(0 \leqq x \leqq 1)$, *and* $x = 0$ *on* $x = 0$, $x = 1$ *on* $x = 1$ $(0 \leqq y \leqq 1)$ *if and only if* T_1 *and* T_2 *are similar.*

This lemma shows that if the smooth annular family (bounded by D_1 and D_2 as above) is cut open along a cross-section pq and mapped on a rectangular family in M so that D_1 corresponds to $y = 0$ and D_2 to $y = 1$ and the direction from $(0, 0)$ to $(1, 0)$ on $y = 0$ to the positive orientation on D_1, then the corresponding identification T is determined up to a class of similar homeomorphisms. Furthermore, if pq is replaced by any other such cross-section $p'q'$, precisely the same class of identifications T is determined. Finally, if two families ϕ', ϕ'' have the same identification class, then there is an o-homeomorphism of ϕ' on ϕ'', taking D'_1 on D''_1, D'_2 on D''_2.

4. 2. Classification of annular families on the sphere. We consider first closed annular families ϕ in $K - N - S$ such that of the two bounding closed curves D_1 and D_2 one, D_1, separates the other, D_2, from N, while D_2 separates D_1 from S. We term two such families ϕ' and ϕ'' equivalent if there is an o-homeomorphism of the family ϕ' on the family ϕ'' taking D'_1 on D''_1 and D'_2 on D''_2.

By the corollary to Theorem 19, with every family ϕ can be associated a sequence of letters, such as

$$bsbbsb, \qquad sbbbsbbsb,$$

corresponding to the successive rings of bent and smooth types. We make the sequence unique for each ϕ by requiring that the letters correspond to the rings, in order of nearness to N, the first letter representing the ring nearest to N. With each of the b's of the sequence is associated a bent annular sub-family of ϕ. This bent annular sub-family has an associated canonical family (*) or (**). We now insert after each b in the sequence the corresponding symbol (*) or (**), so that a sequence might be

$$b(*)sb(**)b(*)sb(**)s.$$

For the smooth sub-families of ϕ the results of 4. 1 permit us to proceed as follows: Suppose the set of all o-homeomorphisms of $0 \leqq y \leqq 1$ on itself subdivided into similarity classes. To each class assign a different abstract index α, so that a collection $A: \alpha, \beta, \cdots$ of indices is formed in one-to-one correspondence with the similarity classes. Then to each smooth annular sub-family ϕ_s corresponds a unique index of A, and two smooth annular sub-families will have the same index if and only if they are equivalent.

We now insert into the already amplified symbol for ϕ the corresponding index of A after each s, so that the symbol might appear thus:

$$b(*)s(\alpha)b(**)b(**)s(\beta)b(*)s(\gamma).$$

This expanded symbol will be denoted by $bs(\phi)$.

In general any finite sequence of b's and s's, with the s's always separated by one or more b's, and with the symbols $(*)$, $(**)$, or (α) (for α in A) inserted as above will be called a finite bs symbol.

THEOREM 21. *For every finite bs symbol σ there exists a closed annular family ϕ on the sphere such that $bs(\phi)$ is σ.*

Proof. Each pair $b(*)$, $b(**)$ or $s(\alpha)$ in the symbol can be realized by a bent or smooth annular family of the corresponding type on the sphere. Equivalent images of these can then be constructed in the correct order and so that successive families have, for example, a latitudinal circle in common. The resulting family is regular by (ii) (§ 1. 2) and has the desired symbol.

THEOREM 22. *Let ϕ' and ϕ'' be two closed annular families on the sphere. Then ϕ' and ϕ'' are equivalent if and only if $bs(\phi') \equiv bs(\phi'')$.*

Proof. The necessity of the condition is an immediate consequence of Theorem 20 and of the fact that the associated canonical family of a bent family is unique.

For the sufficiency it is necessary to set up an o-homeomorphism of ϕ' on ϕ'' taking each smooth or bent sub-family of ϕ' on the corresponding one of ϕ''. This is done by starting with the sub-family (bent or smooth) of ϕ' nearest to N. The fact that the bs symbols are the same shows that there is an o-homeomorphism between this sub-family and the corresponding one of ϕ''. We then proceed step by step to extend the o-homeomorphism to the other sub-families. At each stage, in order to extend the homeomorphism it is necessary to cut open the sub-families to obtain rectangular families (in the case of a bent sub-family two such rectangular families are needed; cf. the proof of Theorem 13). Then by the same reasoning as in the proof of Lemma A of the proof of Theorem 13 the desired extension of the homeomorphism can be made.

All of this discussion applies equally well to the open or half-open annular families ϕ. These are defined as double series $\sum_{-\infty}^{\infty} \phi_i$ or simple series $\sum_{1}^{\infty} \phi_i$ of closed annular families. The families we consider will lie in R and we shall assume that the ϕ_i can be and are so numbered that ϕ_i is nearer than ϕ_{i+1} to N for all i, while ϕ_{i+1} is nearer to S than ϕ_i. Two such families will be termed *equivalent* if there is an o-homeomorphism of one on the other

CALCUTTA CENTRAL LIBRARY UNIVERSITY

preserving the relation of nearness to N on the closed curves. Each ϕ_i can be decomposed as above into a succession of smooth and bent sub-families. The same also holds for each closed annular family $\phi_i + \phi_{i+1} + \cdots + \phi_{i+n}$. It may happen that $\phi_i + \phi_{i+1} + \cdots + \phi_{i+n}$ is itself a smooth annular family, but it could never be a bent family for $n \geqq 1$. It follows that with each half-open family may be associated a sequence of four possible forms:

(h1) $\cdots sssbbsbsbbs$

(h2) $bsbbsbbbsbsbsss \cdots$

(h3) $\cdots bsbbsbbbsb$

(h4) $bsbbsbbbsbsbsbbbsb \cdots$

For forms (h1) and (h3) the range of the index i would be from $-\infty$ to -1; for (h2) and (h4) the range would be from 1 to ∞. For (h1) and (h2) the sequence finally consists wholly of s's if one proceeds far enough to the left or to the right respectively, while for (h3) and (h4) this fails to hold (by definition).

LEMMA. *Let ϕ_h be a half-open annular family on the sphere which can be decomposed into successive annular families all of which are smooth: i. e., ϕ_h corresponds to one of the sequences*

$$\cdots sssss, \qquad sssss \cdots.$$

Then ϕ_h can be mapped homeomorphically on a curve-family $\phi_s[A - D_2]$, where ϕ_s is a smooth annular family filling the annular region A which is bounded by the two closed curves D_1 and D_2 of ϕ_s. Furthermore ϕ_s is unique up to a homeomorphism.

Proof. Map the successive smooth sub-families of ϕ_h homeomorphically on annular families bounded by the successive circles $r = 1$, $r = 3/2$; $r = 3/2$, $r = 7/4$; $r = 7/4$, $r = 15/8$; $\cdots$ in the plane. In each case we choose the image family to be a spiral family (see I, § 1. 8). The result is a family homeomorphic to ϕ_h, and, furthermore, if we adjoin the circle $D_2 : r = 2$, the resulting family ϕ_s is regular and has the desired property.

If two such families are given, then it follows immediately by use of rectangular representations and Theorem 20 that they are homeomorphic.

If we apply the lemma, we see that with each such family ϕ_h may be associated an identification class in a unique way, exactly as with a smooth annular family on the sphere. We can thus express ϕ_h by a symbol $s(\alpha, -\infty)$ or $s(\alpha, +\infty)$, the plus or minus being determined by whether the "missing curve" precedes (minus sign) or succeeds (plus sign) the other closed curves of ϕ_k, where 'precedes' means 'is nearer to N than'.

Accordingly, if we insert the $(*)$, $(**)$ and indices $\alpha, \beta, \gamma, \cdots$ throughout appropriately, the forms (h1) and (h2) reduce to the following types of finite sequences:

(h1) $s(\alpha, -\infty)b(*)b(**)s(\beta)b(**)s(\gamma)b(*)b(*)s(\delta)$

(h2) $b(*)s(\alpha)b(**)b(*)s(\beta)b(*)b(*)b(**)s(\gamma)b(*)s(\epsilon, +\infty)$

The sequence is now unique, and will be denoted by $bs(\phi_k)$. For cases (h3) and (h4) the infinite sequence is unique as it stands, and if the $(*)$, $(**)$, and indices are inserted, it becomes the unique symbol $bs(\phi_h)$.

A similar discussion holds for the open families. There are four forms, illustrated by

(o1) $s(\alpha, -\infty)b(**)b(**)s(\beta)b(**)s(\gamma, +\infty)$

(o2) $s(\alpha, -\infty)b(*)b(**)s(\beta)b(*)s(\gamma)b(*) \cdots$

(o3) $\cdots b(*)s(\alpha)b(*)b(**)s(\beta)b(**)s(\gamma, +\infty)$

(o4) $\cdots b(*)s(\alpha)b(*)b(**)s(\beta)b(*) \cdots$

If any doubly infinite or 'half-infinite' sequence is given, formed by the letters b, s, with each b followed by $(*)$ or $(**)$ and each s followed by an index (α), or, in the case of a first s or a last s, by an index $(\alpha, -\infty)$ or $(\alpha, +\infty)$ respectively, and if the s's are separated by b's, then we term the sequence a *doubly infinite bs symbol* or a *half-infinite bs symbol.*

THEOREM 21a. *To every doubly infinite or half-infinite bs symbol* σ *there exists an open or half-open annular family* ϕ *on the sphere such that* $bs(\phi) \equiv \sigma$.

THEOREM 22a. *Let* ϕ' *and* ϕ'' *be two half-open (or open) annular families on the sphere. Then* ϕ' *and* ϕ'' *are equivalent if and only if* $bs(\phi') \equiv bs(\phi'')$.

(By the identity of two infinite sequences is meant the existence of a one-to-one correspondence between them which preserves the given order and pairs equal elements.)

The proofs of these theorems are essentially the same as the proofs of Theorems 21 and 22.

4. 3. The situation at N and S. We now return to the family F filling R, about which we have made the assumption that it contains one or more closed curves. Consider, as in case (b), the collection $B(N)$ of curves C of F such that $L(C) = N$. The same proof (with simplifications) as for Theorem 10 above ensures that $B(N)$ consists of a flower $Fl(N) = \{C_n\}$ $(n = 1, 2, \cdots)$ plus families filling the sets $\mathfrak{D}(C_n)$.

Case 1. *There is a directed curve C in F, not in $B(N)$, and such that* $L(C, +) = N$. The curve C must then be asymptotic to a closed curve D of F in the negative direction. Choose $\mathfrak{D}(D)$ to include N. Then in $\mathfrak{D}(D) - N$ we have a family in a region homeomorphic to R, and this family is necessarily of the type of case (a). But in an s-family about D the curves of $\mathfrak{D}(D)$ must all spiral asymptotically to D. It follows that in $D + \mathfrak{D}(D) - N$ we have the same situation as in case (b), type $(N\delta, S\epsilon)$, with D playing the rôle of $Fl(N)$ and N the rôle of S. We can choose a closed cross-section γ in $\mathfrak{D}(D)$ and consider, as in § 3.7 and Theorem 16, the double chordal system $DCS(F[\mathfrak{D}(\gamma) - N])$. An existence theorem analogous to Theorem 16 then holds.

Case 2. $B(N) \neq 0$ *and there is no curve C in $F - B(N)$ such that* $L(C, +) = N$. Let C_1 be a curve of $Fl(N)$ and let p_1q be a cross-section from a point p_1 of C_1 to a point q in $F - B(N)$. If p_1q fails to meet a closed curve of F, then there is a closed curve of F nearer to N than all other closed curves of F. Otherwise p_1 is a limit of closed curves of F. In either case the method used in the proof of Theorem 13 is applicable, and we conclude that there is a closed curve D of F such that, if $\mathfrak{D}(D) \supset N$, then $F[D + \mathfrak{D}(D) - \mathfrak{D}Fl(N) - N]$ can be mapped homeomorphically on a curve-family $\phi[A - \Sigma]$, where ϕ is a closed annular family in the plane filling the annulus $A: 1 \leqq r \leqq 2$, and Σ is a non-void closed set on $r = 1$. In other words, $Fl(N)$ acts as one closed curve.

Case 3. $B(N) = 0$ *and there is no curve C in F such that* $L(C, +) = N$. For every closed curve D choose $\mathfrak{D}(D)$ to include N. Let $E = \prod \mathfrak{D}(D)$ for all D of F. Then E is the point N. For if E contains any other point p, then p is on a curve C of F which is not closed, hence is asymptotic to a closed curve in one direction and to N in the other, contrary to assumption. Thus in this case the closed curves D close down on N as limit.

All of this discussion holds similarly for S.

4.4. Classification in case (c). We can indicate the combinations at N and S by the symbols (Ni, Sj) $(i = 1, 2, 3;\ j = 1, 2, 3)$. The combination (Ni, Sj) will be regarded as the same as (Nj, Si), since a renaming of N and S makes them the same. There are then in all six possible types.

Consider the types (Ni, Sj) with $i = 1, 2$, $j = 1, 2$. Suppose first that F contains at least two closed curves. Then in all four of these cases F can be mapped homeomorphically on a family F^0 consisting of a closed annular family ϕ^0, bounded by two latitudinal circles $D_N{}^0$ and $D_S{}^0$, from which a closed set on either or both bounding circles may have been removed, plus families in $\mathfrak{D}(D_N{}^0)$ and $\mathfrak{D}(D_S{}^0)$, where $\mathfrak{D}(D_N{}^0) \supset N$ and $\mathfrak{D}(D_S{}^0) \supset S$. The structures in $\mathfrak{D}(D_N{}^0)$ and $\mathfrak{D}(D_S{}^0)$ correspond to the cases 1 and 2 at N and S. If case 1 holds at N, then we have in $D_N{}^0 + \mathfrak{D}(D_N{}^0)$ the type of family described under case 1 as filling $D + \mathfrak{D}(D)$. If case 2 holds at N, the structure in $\mathfrak{D}(D_N{}^0)$ is that of $F[\mathfrak{D}Fl(N)]$, as described under case 2, while $Fl(N)$ is represented by open arcs on $D_N{}^0$. The closed annular family ϕ^0 bounded by $D_N{}^0$ and $D_S{}^0$ is completely characterized by the discussion above: i.e., the symbol $bs(\phi^0)$ gives all the information. If we use the facts previously obtained for homeomorphism of two families, we conclude that *two curve families F_1 and F_2 of these types can be homeomorphic if and only if forms $F_1{}^0$ and $F_2{}^0$ of each can be chosen such that each consists of an annular family $\phi_i{}^0$ $(i = 1, 2)$ (with possible deletions on the boundaries) and families in $\mathfrak{D}(D^0{}_{iN})$ and $\mathfrak{D}(D^0{}_{iS})$, whereby $bs(\phi^0{}_1) \equiv bs(\phi^0{}_2)$, and the families $F^0{}_1[D^0{}_{1N} + \mathfrak{D}(D^0{}_{1N})]$ and $F^0{}_1[D^0{}_{1S} + \mathfrak{D}(D^0{}_{1S})]$ have respectively the same structure as $F^0{}_2[D^0{}_{2N} + \mathfrak{D}(D^0{}_{2N})]$ and $F^0{}_2[D^0{}_{2S} + \mathfrak{D}(D^0{}_{2S})]$.* These last structures have already been characterized by Theorems 17 and 18 above.

If F contains just one closed curve, the same discussion holds except for the type $(N1, S1)$. Here the annulus collapses to a single closed curve D and we have only to consider a family in $\mathfrak{D}(D)$ and a family in $\mathfrak{D}^*(D)$ of type 1. The criterion thus holds, with $\phi_i{}^*$ replaced by a single closed curve.

Consider next the types $(N3, Sj)$ with $j = 1$ or 2. Here we can map F homeomorphically on a family F^0 consisting of a half-open annular family bounded by the equator D and lying in $\mathfrak{D}(D)$, where $\mathfrak{D}(D) \supset N$, plus a family in $\mathfrak{D}^*(D)$ of the types 1 and 2. A similar criterion can then be set up for homeomorphism.

Finally consider the type $(N3, S3)$. Here F is an open annular family to which Theorems 21a and 22a apply.

It is to be remarked that it is only in these last two cases that an infinite number of bent annular sub-families can appear in F.

PART B. Isolated Singularities of a Family on a Surface.

1. **Reduction to a family of *R*.** Let Φ be a surface, which may be open or closed. Let ψ be an open sub-region on Φ, and let F be a regular curve-family filling ψ. Let p be an isolated point of $\Phi - \psi$, and let $U(p)$

be an open subset of $\psi + p$ containing p. The problem is to determine all the possible structures which F may have in $U(p) - p$.

This can be solved if we assume only that $U(p)$ is simply-connected. There certainly exist such $U(p)$ and they can be chosen arbitrarily small. Since the size of $U(p)$ will be of no import for the procedure below, the most information will be obtained if $U(p)$ is chosen to be 'as large as possible,' while remaining simply-connected. For example, if F fills the surface of a torus with the exception of one point p, we choose two generating circles not meeting p and choose $U(p)$ as the complementary region of the two circles.

Since $U(p)$ is simply-connected, $U(p)$ can be mapped homeomorphically on $K - S$ so that p has N as image. The part of F in $U(p) - p$ thus becomes a regular curve-family F' filling R. Hence the problem of the structure of F in $U(p) - p$ is completely solved by our results above for the family filling R.

2. **Local considerations.** It may be asked what is the structure of F in an arbitrarily small set $U(p)$. This means of course sacrificing a good part of the information which could be obtained by taking $U(p)$ large. But the question is worth considering, especially since in the study of differential equations it is imporant to know the possible structures of the solutions in an arbitrarily small neighborhood of a singular point. (See Fischer [2], Forster [3], Frommer [4], Lohn [10], v. Mises [11], and Petrovitch [12]).

We can first take $U(p)$ simply-connected and map on the sphere as above. The question then becomes one of determining the structure of the family F' in the neighborhood of N. Suppose that F' contains a curve C such that $L(C, +) = N$. Then for a sufficiently small neighborhood V of N, C will meet the boundary of V. It follows that in V, F' has the structure of case (a). If F' contains no such curve C, then in V only case (c) can hold, and in fact N must be the limit of closed curves of F' in V. V can be chosen as the interior $\mathfrak{D}(C)$ of such a closed curve C, and it follows that in $V - N$, F' is a half-open annular family.

THEOREM 23. *The only possible structures of a regular curve-family on a surface in the (arbitrarily small) neighborhood of an isolated singularity p are those of case (a) or those of a half-open annular family with p as limit.*

It can further be shown that every regular curve-family filling R is homeomorphic to a differentiable family (see Kaplan [7]). Hence the above structures at p are precisely those which arise from differential equations.

BIBLIOGRAPHY.

Kaplan [5] gives a list of papers on the topological theory of curve-families. Petrovitch [12] gives an extended bibliography on the subject, up till 1931, from the point of view of differential equations. The present list is intended only as a supplement to these.

1. A. Andronov and L. Pontrjagin, "Systèmes grossiers," *C. R. Acad. Sci. URSS*, N. S., vol. 14 (1937), pp. 247-250.
2. H. J. Fischer, "Der Verlauf von Integralkurven einer Differentialgleichung erster Ordnung," *Deutsche Math.*, vol. 3 (1938), pp. 152-188.
3. H. Forster, "Über das Verhalten der Integralkurven einer gewöhnlichen Differentialgleichung erster Ordnung in der Umgebung eines singulären Punktes," *Math. Zeits.*, vol. 43 (1937), pp. 271-320.
4. M. Frommer, "Über das Auftreten von Wirbeln und Strudeln in der Umgebung rationaler Unbestimmtheitsstellen," *Mathematische Annalen*, vol. 109 (1934), pp. 395-424.
5. W. Kaplan, "Regular Curve-families Filling the Plane," I, *Duke Mathematical Journal*, vol. 7 (1940), pp. 154-185.
6. ———, "Regular Curve-families Filling the Plane," II, *Duke Mathematical Journal*, vol. 8 (1941), pp. 11-46.
7. ———, "Differentiability of Regular Curve-families on the Sphere," Lectures on Topology (Ann Arbor, 1941), pp. 299-301.
8. B. v. Kerékjártó, "On a Geometrical Theory of Continuous Groups," I, *Annals of Mathematics*, ser. 2, vol. 27 (1925), pp. 105-117.
9. E. Leontovič and A. Mayer, "Sur les trajectoires qui déterminent la structure qualitative de la division de la sphère en trajectoires," *C. R. Acad. Sci. URSS*, N. S., vol. 14 (1937), pp. 251-254.
10. E. R. Lohn, "Über singuläre Punkte gewöhnlicher Differentialgleichungen," *Math. Zeits.*, vol. 44 (1938), pp. 507-530.
11. R. v. Mises, "Über den Verlauf der Integralkurven einer Differentialgleichung erster Ordnung," *Compositio Math.*, vol. 6 (1939), pp. 203-220.
12. M. Petrovitch, "Intégration qualitative des équations différentielles," *Mém. des Sc. Math.*, Fasc. 48 (Paris, 1931).

A CONTRIBUTION TO THE FOUR COLOR PROBLEM.*

By Chaim Chojnacki.

(1, 1) In the study of the four color problem we assume[1] a map divided by a connected trihedral network into a finite number of polygons.

A map A is *reducible* if it is possible to form a map B containing a smaller number of polygons than A so, that from the assumption that the map B is colorable it follows that the map A is colorable too. In such cases we shall say that the map A is *reduced* to the map B.

If there are maps which are not colorable, there exists among them a map M which contains the least number of polygons. Such a map is called *minimal.* This definition implies that a minimal map is not reducible.

(1, 2) A sequence of polygons $P^1, P^2, \cdots, P^n$ will be called a *chain* if every polygon P^i touches P^{i+1} $(i = 1, 2, \cdots, n-1)$.

A closed chain (P^n touches P^1) which divides the map into at least two distinct parts will be called a *ring.*

A well-known theorem of Birkhoff[2] states:

(1, 2, 1) "A minimal map cannot contain a ring composed of less than five polygons, or a ring of five polygons enclosing more than one polygon."

Franklin[3] proved the following theorem:

(1, 2, 2) "A minimal map does not contain a polygon of n sides touching $n-1$ or n pentagons."

(1, 3) It will be proved in this paper for polygons having n $(6 \leq n \leq 9)$ sides, that such a polygon in a minimal map cannot touch more than $n-4$ successive pentagons.

This result will enable us to sharpen Reynolds'[4] theorem stating that a

---

* Received March 3, 1939.

[1] For a general account of the subject see M. A. Sainte-Laguë, "Géometrie de situation et jeux," *Mémorial des Sciences Mathematiques XLI,* and the thesis of Alfred Errera, *Du coloriage des cartes et de quelques questions d'analysis situs,* Bruxelles 1921.

[2] George D. Birkhoff, "The reducibility of maps," *American Journal of Mathematics,* vol. 35 (1913), pp. 115-128.

[3] Philip Franklin, "The four color problem," *American Journal of Mathematics,* vol. 44 (1922), pp. 225-236.

[4] C. N. Reynolds, Jr., "On the problem of coloring maps in four colors II," *Annals of Mathematics,* vol. 28 (1927), p. 492.

minimal map must include either a hexagon, or a heptagon, or an octagon. We shall prove that a minimal map must contain a hexagon or a heptagon. We shall even prove that in a minimal map

$$3p_6 + p_7 \geq p_9 + \sum_{i>9} (i-9)p_i + 24,$$

where p_i denotes the number of i-gons in the map.

(2) THEOREM: *In a minimal map there is no polygon containing* n $(6 \leq n \leq 9)$ *sides touching* $n-3$ *successive pentagons.*

Proof: Let us denote pentagons by 5, hexagons by 6, arbitrary polygons (which may be pentagons or hexagons) by an asterisk *. We have to consider the following four configurations

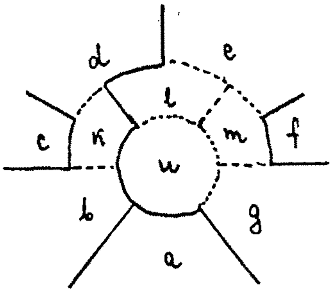
(2, 1) 555*** (a hexagon surrounded by three pentagons and any three polygons in this order)

(2, 2) 5555***

(2, 3) 55555***

(2, 4) 555555***.

(2, 1) 555***[5]



In this case the hexagon u is surrounded by three pentagons k, l, m, and three arbitrary polygons a, b, g. Let us consider the configuration K which contains also the polygons c, d, e, f touching the former, as in the figure. We denote $k+l+m+u$ by s and $a+b+c+d+e+f+g$ by t.

(2, 1, 1) Assume that the minimal map A contains K. According to the theorem of Birkhoff (1, 2, 1) the ploygons of t are all different, and they have no common sides except those which are marked on the figure.

The map A will be reduced by cancelling the dotted boundaries. So the polygons b, k, d will coalesce to one polygon, and the polygons e, l, m, u, g to another and so we shall get a map B which contains 6 polygons less than the map A. According to our assumption the map B is colorable.

---

[5] *Loc. cit.* [4].

Let the color of the polygon a be denoted by 1, the color of the polygon $b + k + d$ by 2, and the color of $e + l + m + u + g$ by 3. In short we shall write

$$a = 1, \quad b = d = 2, \quad e = g = 3$$

The polygons c, f may have the following colors:

	c	f
(1)	1	1
(2)	1	2
(3)	1	4
(4)	3	1
(5)	3	2
(6)	3	4
(7)	4	1
(8)	4	2
(9)	4	4

Let us return to the map A and leave to all polygons outside s the same colors as in B. This is allowed according to (2, 1, 1). Presently we shall see that some cases among (1)-(9) are identical with or symmetrical to one another by change of the colors, viz. case (4) is symmetrical to the case (2), (7) to (3), and (8) to (6).

There remain finally the following cases:

	c	f
(1)	1	1
(2)	1	2
(3)	1	4
(5)	3	2
(6)	3	4
(9)	4	4

In some of these cases it is possible to color the polygons of s directly as shown below, in other cases we shall use a special procedure which we shall perform in every case in a corresponding table L as marked below (the tables L are denoted by Roman figures).

(2, 1, 2)

c	f	k	l	m	u	L
1	1	3	1	2	4	
1	2					II
1	4	3	1	2	4	
3	2					III
3	4					I
4	4	3	1	2	4	

We begin with the composition of table I.

The following coloration of t is given

a	b	c	d	e	f	g
1	2	3	2	3	4	3

Consider the union $\overset{13}{c}$ of all the chains of polygons colored 1 and 3 in A-s containing the polygon c. We write

	a	b	c	d	e	f	g
	1	2	3	2	3	4	3
$\overset{13}{c}$							

In this case three possibilities can be distinguished:

ce: the polygon e is included in $\overset{13}{c}$;

ca: the polygon a (and g) are included in $\overset{13}{c}$;

$c/$ ("c only") no polygon of t except c is included in $\overset{13}{c}$.

In the case ce $\overset{24}{d}$ contains no polygon of t except d; so that the interchange of the colors 2 and 4 in $\overset{24}{d}$ will not cause any change in the coloring of t besides the change of the color of d. So we get the coloration of the polygons of t enabling us to assign the following coloration of s:

	a	b	c	d	e	f	g	k	l	m	u
ce	1	2	3	4	3	4	3	1	2	1	4

In the case ca we shall get similarly

	a	b	c	d	e	f	g	k	l	m	u
ca	1	4	3	2	3	4	3	1	4	1	2

In the case $c/$ the interchange of the colors 1 and 3 in $\overset{13}{c}$ will not cause any change in the coloration of t besides the change of the color of c. We obtain

	a	b	c	d	e	f	g	k	l	m	u
c/	1	2	1	2	3	4	3	3	1	2	4

It follows that in case $c = 3$, $f = 4$ the map A is colorable. The above considerations may be presented in the following abridged form:

I	a	b	c	d	e	f	g	k	l	m	u
	1	2	3	2	3	4	3				
$\overset{13}{c}e$	1	2	3	4	3	4	3	1	2	1	4
ca	1	4	3	2	3	4	3	1	4	1	2
$c/$	1	2	1	2	3	4	3	3	1	2	4

The tables II and III will be formed similarly

II	a	b	c	d	e	f	g	k	l	m	u
	1	2	1	2	3	2	3				
$\overset{14}{c}a$	1	3	1	2	3	2	3	4	1	4	2
$c/$	1	2	4	2	3	2	3	symmetrical to I			

and thus colorable as shown before.

III.	a	b	c	d	e	f	g	k	l	m	u
	1	2	3	2	3	2	3				
$\overset{12}{f}d$	1	2	3	2	4	2	3	1	3	1	4

$fa/$ (this sign denotes that the polygon a is included in $\overset{12}{f}$ but no other polygon of t besides f is included in $\overset{12}{f}$. In this case $\overset{34}{g}$ does not contain any polygon of t except g, and $\overset{12}{f}$ includes only the polygons f, a (and b) of t, so that the interchange of the colors 3 and 4 in $\overset{34}{g}$ and the colors 1 and 2 in $\overset{12}{f}$ will not cause any changes in the coloring of t besides the change of the colors of g, f, a, and b)

	2	1	3	2	4	1	4	4	1	2	3
$f/$	1	2	3	2	4	1	3	1	3	2	4

The explanations given in (2, 1) will not be repeated henceforth in analogous places. Henceforth only the following notes are given:

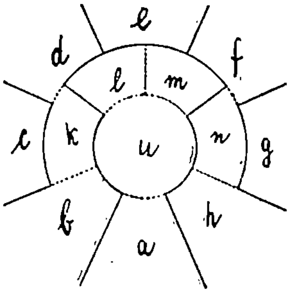
To every configuration K a figure will be added with dotted sides the cancelling of which reduces the given map.

To the figure will be added a list of the colorations of K-s which correspond to this reduction, like the table (2, 1, 2). Here s denotes the sum of all those polygons of K which touch polygons of K only. In this table only one case from every group of cases identical (by a permutation) or symmetrical will be included.

To every table of this kind will be attached the tables L belonging to those colorations of K-s which do not immediately permit the coloration of s. The tables L will be given in the form of the tables I, II and III as before.

It should be further noted with regard to the configurations K, that all polygons of K are different, and have no common sides besides those denoted on the corresponding figure, as is easily shown (for any K to appear) by Birkhoff's theorem, cf. (2, 1, 1).

(2, 2) 5555***



$b = d = 2,\ a = 1,\ e = 3,\ f = h.$

$f = h$	c	g	k	l	m	n	u	L
2	1	1						*I*
2	1	3	4	1	4	1	3	
2	1	4	4	1	4	1	3	
2	3	3	4	1	4	1	3	
2	3	4	4	1	4	1	3	
2	4	4						*II*
4	1	1						*III*
4	1	2	4	1	2	1	3	
4	1	3	4	1	2	1	3	
4	3	2	4	1	2	1	3	
4	3	3	4	1	2	1	3	
4	4	2	1	4	2	1	3	

I

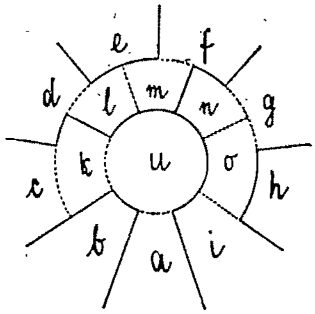
	a	b	c	d	e	f	g	h	k	l	m	n	u
	1	2	1	2	3	2	1	2					
24 *fh*	1	2	1	2	3	2	3	2	4	1	4	1	3
fb/	1	2	1	4	3	2	1	2	3	2	1	3	4
fd/	1	2	1	4	3	4	1	2	3	2	1	3	4
fdb	1	2	3	2	3	2	1	2	1	4	1	4	3
f/	1	2	1	2	3	4	1	2	3	1	2	3	4

II

	a	b	c	d	e	f	g	h	k	l	m	n	u
	1	2	4	2	3	2	4	2					
12 *fd*	1	2	4	2	4	2	4	2	1	3	1	3	4
fh	1	2	4	2	3	2	3	2	1	4	1	4	3
f/	1	2	4	2	3	1	4	2	3	1	2	3	4

III	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*k*	*l*	*m*	*n*	*u*
	1	2	1	2	3	4	1	4					
$\overset{24}{bd}$	1	2	3	2	3	4	1	4	1	4	1	2	3
bh	3	2	1	2	3	4	1	4	3	4	2	3	1
bf/	1	4	1	2	3	2	1	4	3	1	4	3	2
b/	1	4	1	2	3	4	1	4	3	4	1	3	2

(2, 3) 55555***



$a = c = 1, d = f = 2, g = i = 3.$

b	*e*	*h*	*k*	*l*	*m*	*n*	*o*	*u*	*L*
2	1	1							*III*
2	1	2							*IV*
2	1	4							*XI*
2	3	1							*VIII*
2	3	2							*VII*
2	3	4							*VI*
2	4	1	3	1	3	1	2	4	
2	4	2							*IX*
2	4	4	3	1	3	1	2	4	
3	1	1	4	3	4	1	4	2	
3	1	2	4	3	4	1	4	2	
3	1	4							*V*
3	3	1	4	1	4	1	4	2	
3	3	2	4	1	4	1	4	2	
3	3	4							*X*
3	4	1	4	1	3	1	4	2	
3	4	2	4	1	3	1	4	2	
3	4	4	4	1	3	4	1	2	
4	1	1	3	4	3	1	4	2	
4	1	2	3	4	3	1	4	2	
4	1	4	3	4	3	4	1	2	
4	3	1	3	1	4	1	4	2	
4	3	2	3	1	4	1	4	2	
4	3	4	3	4	1	4	1	2	
4	4	1	3	1	3	1	4	2	
4	4	2	3	1	3	1	4	2	
4	4	4	3	1	3	4	1	2	

I	*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
	2 3 1	3 1 3	2 4 3			
13 *ib*	4 3 1	3 1 3	2 4 3			
14 *ha*	4 3 1	3 1 3	2 4 2	2 4 2	4 3	1
he	4 3 1	3 1 2	3 4 3	2 4 3	4 2	1
h/	4 3 1	3 1 3	2 1 3			
34 *fd*	4 3 1	3 2 3	2 1 3	4 1 4	1 4	2
fi	4 3 1	3 1 3	1 2 3	4 2 4	2 4	1
f/	4 3 1	3 1 4	2 1 3	2 4 2	3 4	1
hc	4 2 1	3 1 3	2 4 2			
28 *db*	4 2 4	3 1 3	2 4 2	1 4 2	4 1	3
df	4 2 1	3 4 3	2 4 2	4 1 2	4 1	3
di	1 2 4	3 1 3	2 4 2	1 4 2	4 1	3
d/	4 2 1	2 1 3	2 4 2	4 3 2	4 3	1
i/	2 3 1	3 1 3	2 4 1			
28 *fd*	2 3 1	3 4 3	2 4 1	2 1 2	1 3	4
fa	2 3 1	3 1 3	2 1 4	2 4 2	4 3	1
f/	2 3 1	3 1 2	3 4 1			
24 *ah*	2 3 1	3 1 2	3 4 3	2 4 3	4 2	1
af	2 1 3	1 3 2	3 4 1	2 4 1	4 2	3
a/	4 3 1	3 1 2	3 4 1			
23 *b/*	4 2 1	3 1 2	3 4 1	4 2 4	1 2	3
bd	4 3 4	3 1 2	3 4 1			
13 *ib*	2 3 4	3 1 2	3 4 1	1 2 3	1 2	4
id	2 3 2	3 1 2	3 4 1	1 2 3	1 2	4
ig	4 3 4	3 1 2	3 2 1	1 4 3	1 4	2
i/	4 3 4	3 1 2	3 4 3	2 4 3	4 2	1
bg	1 3 1	3 1 2	3 1 4			
12 *a/*	2 3 1	3 1 2	3 1 4	2 4 3	4 2	1
ac	1 4 1	3 1 2	3 1 4			
18 *da*	1 2 1	3 1 2	3 1 4	4 2 4	1 2	3
dg	1 4 1	3 1 4	3 1 4	2 4 2	1 2	3
d/	1 4 3	1 3 2	3 1 4	2 4 1	4 2	3
ah	1 3 1	3 1 2	3 1 3			
28 *db*	1 3 4	3 1 2	3 1 3	1 4 3	1 4	2
di	1 3 1	3 4 2	3 4 3	4 1 3	4 1	2
df	1 3 1	3 4 2	3 1 3	2 1 3	1 2	4
d/	1 3 1	2 1 2	3 1 3	4 3 4	1 4	2
ae	1 4 1	4 1 2	3 1 4			

I	*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
^{34} *db*	1 4 2	4 1 2	3 1 4	1 2 4	1 2	3
di	2 4 2	4 1 2	3 1 4	1 2 4	1 2	3
d/	1 4 1	3 1 2	3 1 4			
^{13} *da*	1 2 1	3 1 2	3 1 4	4 2 4	1 2	3
dg	1 4 1	3 1 4	3 1 4	2 4 2	1 2	3
d/	1 4 3	1 3 2	3 1 4	2 4 1	4 2	3
dg	1 4 1	4 2 1	3 1 4			
^{12} *h/*	1 4 1	4 2 1	3 2 4	3 1 3	4 1	2
hc/	1 4 2	4 2 1	3 2 4	3 1 3	4 1	2
ha	1 4 1	4 2 1	3 1 3			
^{14} *hf*	1 4 1	4 2 1	2 1 3	3 1 4	3 4	2
ha	1 4 1	4 2 1	3 1 2	2 1 4	2 4	3
h/	1 4 1	4 2 1	3 4 3	3 1 3	4 1	2
hf	1 4 1	4 2 1	4 1 4			
^{24} *gi*	1 4 1	4 2 1	4 3 4	3 1 4	3 1	2
ge	1 4 1	4 2 3	4 1 4	3 1 4	1 3	2
gb	3 4 1	4 2 1	4 3 4	3 1 4	3 1	2
g/	1 4 1	4 2 1	2 1 4	3 1 3	4 3	2

II	*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
	1 3 2	1 3 2	1 4 3			
^{12} *a/*	2 3 2	1 3 2	1 4 3	4 2 4	3 2	1
ac	1 4 2	1 3 2	1 4 3			
^{13} *ad*	1 2 4	1 3 2	1 4 3	3 2 1	3 2	4
ag	1 4 2	1 3 2	1 2 3	3 4 1	3 4	2
a/	3 4 2	1 3 2	1 4 1	3 4 1	4 3	2
ag	1 3 2	1 3 2	1 3 4			
^{23} *b/*	1 2 3	1 3 2	1 3 4	4 2 1	4 2	3
bh	4 3 2	1 3 2	1 3 1			
^{34} *ae*	4 3 1	2 3 2	1 3 1	4 1 4	3 4	2
ah	4 3 2	1 3 2	1 3 2	4 2 4	3 4	1
a/	3 4 2	1 3 2	1 3 1	3 4 1	3 4	2
be	1 3 2	4 3 2	1 3 4			
^{12} *ac*	1 4 2	4 3 2	1 3 4	1 2 1	4 2	3
ag	1 3 2	4 3 2	1 4 3	1 2 1	3 2	4
a/	2 3 2	4 3 2	1 3 4			
^{23} *ce*	2 3 2	1 3 2	1 3 4	4 2 4	3 2	1
ch	2 3 2	4 3 2	1 3 1	1 2 1	3 2	4
c/	3 2 3	4 3 2	1 3 4			

II			*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
	24	*bd*	3 2 1	4 3 2	1 3 4	3 2 4	3 2	1
		bf	3 2 1	4 1 2	1 3 4	3 2 4	3 2	1
		bi	1 2 3	4 3 2	1 3 4	1 2 1	4 2	3
		b/	3 4 3	4 3 2	1 3 4			
	14	*db*	3 4 2	4 3 2	1 3 4	3 2 4	3 2	1
		di	2 4 2	4 3 2	1 3 4	3 2 4	3 2	1
		d/	3 4 3	1 3 2	1 3 4			
	13	*da*	3 2 3	1 3 2	1 3 4	4 2 4	3 2	1
		dg	3 4 3	1 3 4	1 3 4	2 4 2	3 2	1
		d/	3 4 1	3 1 2	1 3 4	2 4 3	4 2	1
		dg	3 4 3	4 2 3	1 3 4			
	23	*h/*	3 4 3	4 2 3	1 2 4	1 3 1	4 3	2
		hc/	3 4 2	4 2 3	1 2 4	1 3 1	4 3	2
		ha	3 4 3	4 2 3	1 3 1			
	34	*ha*	3 4 3	4 2 3	1 3 2	2 3 4	2 4	1
		hf	3 4 3	4 2 3	2 3 1	1 3 4	1 4	2
		h/	3 4 3	4 2 3	1 4 1	1 3 1	4 3	2
		hf	3 4 3	4 2 3	4 3 4			
	24	*gi*	3 4 3	4 2 3	4 1 4	2 3 4	2 3	1
		ge	3 4 3	4 2 1	4 3 4	2 3 4	3 2	1
		gb	3 4 1	4 2 1	4 3 4	2 3 4	3 2	1
		g/	3 4 3	4 2 3	2 3 4	1 3 1	4 1	2

III			*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
			1 2 1	2 1 2	3 1 3			
	23	*df*	1 2 1	2 4 2	3 1 3	3 1 3	1 2	4
		di	4 2 4	2 1 2	3 1 3	3 4 3	4 2	1
		db/	1 3 4	3 1 2	3 1 3	1 2 3	1 2	4
		d/	1 2 1	3 1 2	3 1 3			
	23	*bd*	1 2 4	3 1 2	3 1 3	1 2 3	1 2	4
		bf	1 2 4	3 4 2	3 1 3	1 2 3	1 2	4
		bi	4 2 1	3 1 2	3 1 3	4 2 3	4 2	1
		b/	1 3 1	3 1 2	3 1 3			
	23	*df*	1 3 1	3 4 2	3 1 3	4 1 3	1 4	2
		db	1 3 4	3 1 2	3 1 3	1 4 3	1 4	2
		di	4 3 4	3 1 2	3 1 3	1 4 3	1 4	2
		d/	1 3 1	2 1 2	3 1 3	4 3 4	1 4	2

IV	*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
	1 2 1	2 1 2	3 2 3			
^{14} *ac*	1 3 1	2 1 2	3 2 3	4 3 4	1 4	2
ae	1 2 1	2 1 3	2 3 2			
^{23} *db*	1 2 4	2 1 3	2 3 2	1 4 2	1 4	3
df	1 2 1	2 4 3	2 3 2	4 1 2	1 4	3
d/	1 2 1	3 1 3	2 3 2	4 2 4	1 4	3
a/	4 2 1	2 1 2	3 2 3			
^{13} *gi*	4 2 1	2 1 2	3 4 3	3 4 3	4 2	1
ge	4 2 1	2 1 4	3 2 3	3 4 3	2 4	1
gc	4 2 1	4 1 4	3 2 3	3 2 3	2 4	1
g/	4 2 1	2 1 2	1 2 3	4 3 4	3 4	1

V	*a b c*	*d e f*	*g h i*	*k l m*	*n o*	*u*
	1 3 1	2 1 2	3 4 3			
^{12} *ac*	1 4 1	2 1 2	3 4 3	3 4 3	4 1	2
a/	2 3 1	2 1 2	3 4 3			
^{14} *hc*	3 2 1	2 1 2	3 4 2	identical with I		
h/	2 3 1	2 1 2	3 1 3			
^{12} *ca*	2 4 1	2 1 2	3 1 3	3 4 3	4 2	1
ch	2 3 1	2 1 2	4 1 3	4 3 4	3 2	1
c/	2 3 2	1 2 1	3 1 3	4 3 4	2 4	1
he	2 3 1	2 1 3	2 4 3			
^{23} *d/*	2 3 1	3 1 3	2 4 3	identical with I		
df	2 3 1	2 4 3	2 4 3	symmetrical to II		
db	2 3 4	2 1 3	2 4 3			
^{34} *f/*	2 3 4	2 1 4	2 4 3	1 3 2	3 1	4
fh	2 3 4	2 1 3	1 4 3			
^{23} *ad*	2 3 1	2 1 3	1 4 3	4 3 2	4 2	1
af	2 3 4	2 1 3	4 1 3	1 3 2	1 2	4
a/	3 2 4	2 1 3	1 4 2	3 4 2	4 3	1
fc	2 3 4	1 2 3	2 4 3			
^{12} *ge*	2 3 4	1 2 4	2 4 3	2 3 1	3 1	4
g/	2 3 4	1 2 3	1 4 3			
^{13} *i/*	2 3 4	1 2 3	1 4 1	2 3 1	2 3	4
ig	2 3 4	1 2 3	1 2 3	2 3 4	2 4	1
ib	4 3 4	1 2 3	1 4 3			

V

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*k*	*l*	*m*	*n*	*o*	*u*
24															
he	4	3	4	1	2	1	3	4	3	2	4	3	4	2	1
hc	4	1	4	1	2	3	1	4	1	2	4	1	4	2	3
ha/	2	3	4	1	2	3	1	2	3	2	3	4	2	4	1
h/	4	3	4	1	2	3	1	2	3	2	3	4	2	4	1
id	4	3	2	1	2	3	1	4	3						
24															
ha	4	3	2	1	2	3	1	4	1	4	3	1	4	3	2
hc	4	1	2	1	2	3	1	4	1	4	3	1	4	3	2
he/	4	3	2	1	4	1	3	2	3	4	2	3	2	4	1
h/	4	3	2	1	2	3	1	2	3	4	3	4	2	4	1
ga	2	3	4	1	2	3	2	3	4						
24															
ac	2	1	4	1	2	3	2	3	4	2	4	1	4	1	3
ae	2	1	4	3	2	3	2	3	4	2	4	1	4	1	3
ag/	4	3	4	1	2	3	4	1	2	2	3	4	2	3	1
a/	4	3	4	1	2	3	2	3	2						
12															
g/	4	3	4	1	2	3	1	3	2	2	3	4	2	4	1
gi	4	3	4	1	2	3	2	4	2						
24															
ac	4	1	4	1	2	3	2	4	2	2	4	1	4	1	3
ae	4	3	4	1	2	1	2	4	2	2	4	3	4	3	1
a/	2	3	4	1	2	3	4	2	4	2	3	4	2	3	1
ge	4	3	4	1	2	4	2	3	2						
14															
fd	4	3	4	1	3	4	2	3	2	2	4	2	3	4	1
fa	4	3	4	1	2	4	3	2	3	2	4	3	2	4	1
f/	4	3	4	1	2	1	2	3	2	2	3	4	3	4	1

VI

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*k*	*l*	*m*	*n*	*o*	*u*
	1	2	1	2	3	2	3	4	3						
12															
fd	1	2	1	2	4	2	3	4	3	3	1	3	1	2	4
f/	1	2	1	2	3	1	3	4	3						
28															
ig	1	2	1	2	3	1	3	1	3	symmetrical to III					
ib	4	2	1	2	3	1	3	4	3	3	4	2	4	2	1
id	4	2	4	2	3	1	3	4	3	3	4	2	4	2	1
i/	1	2	1	2	3	1	3	4	2	4	1	2	4	1	3

VII

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*k*	*l*	*m*	*n*	*o*	*u*
	1	2	1	2	3	2	3	2	3						
14															
ca	1	3	1	2	3	2	3	2	3	4	1	4	1	4	2
c/	1	2	4	2	3	2	3	2	3						

VII		a b c	d e f	g h i	k l m	n o	u
	12						
	h/	1 2 4	2 3 2	3 1 3	symmetrical to VI		
	ha	1 2 4	2 3 2	3 2 4	1 4 1	4 1	3
	hd	1 2 4	2 4 2	4 2 3	1 3 1	3 1	4
	hf	1 2 4	2 3 2	4 2 3			
	34						
	ec	1 2 4	1 3 2	4 2 3	3 2 1	3 1	4
	eg	1 2 4	2 3 1	4 2 3	3 1 2	3 1	4
	ei	1 2 4	2 3 1	4 1 3	3 1 2	3 2	4
	e/	1 2 4	2 4 2	4 2 3	1 3 1	3 1	4

VIII		a b c	d e f	g h i	k l m	n o	u
		1 2 1	2 3 2	3 1 3			
	12						
	f/	1 2 1	2 3 1	3 1 3	symmetrical to III		
	fd	1 2 1	2 4 2	3 1 3	3 1 3	1 2	4
	fh	1 2 1	2 3 2	4 1 3			
	12						
	df	1 2 1	2 4 2	4 1 3	3 1 3	1 2	4
	dh	1 2 1	2 3 2	4 1 4	4 1 4	1 2	3
	d/	2 1 2	1 3 2	4 1 3	3 2 1	3 2	4

IX		a b c	d e f	g h i	k l m	n o	u
		1 2 1	2 4 2	3 2 3			
	34						
	e/	1 2 1	2 3 2	3 2 3	identical with VII		
	eg	1 2 1	2 4 1	3 2 3	3 1 3	2 1	4
	ei	2 1 2	1 4 2	3 2 3			
	12						
	hf	2 1 2	1 4 2	4 2 3	3 2 1	3 1	4
	ha	2 1 2	1 4 2	3 2 4	4 2 1	4 1	3
	h/	2 1 2	1 4 2	3 1 3	3 2 3	1 2	4

X		a b c	d e f	g h i	k l m	n o	u
		1 3 1	2 3 2	3 4 3			
	13						
	gi	1 3 1	2 3 2	3 2 3	4 1 4	1 4	2
	g/	1 3 1	2 3 2	1 4 3			
	34						
	eb	1 3 2	1 3 2	1 4 3	identical with II		
	eh	1 3 1	2 3 1	2 4 3	4 1 4	3 1	2
	e/	1 3 1	2 4 2	1 4 3			
	13						
	gi	1 3 1	2 4 2	1 2 3	4 3 1	3 4	2
	g/	1 3 1	2 4 2	3 4 3	4 3 1	4 1	2
	ge	1 3 1	2 3 4	3 4 3			
	14						
	ca	1 2 1	2 3 4	3 4 3	3 1 2	1 2	4
	cf	1 3 1	3 2 4	3 4 3	2 1 3	1 2	4
	c/	1 3 4	2 3 4	3 4 3			

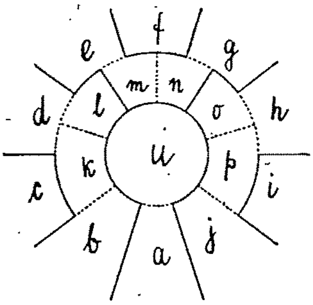
X

	a b c	*d e f*	*g h i*	*k l m*	*n o*	*u*
18						
eg	1 3 4	2 3 2	3 4 3	1 4 1	4 1	2
eb	1 3 2	4 3 4	3 4 3	1 2 1	2 1	4
e/	1 3 4	2 1 4	3 4 3	1 3 2	1 2	4
ch	1 2 1	2 3 4	3 4 2			
23						
ge	1 2 1	2 3 1	3 4 2	4 1 2	4 1	3
gi	1 2 1	2 3 4	3 1 2	4 1 2	1 4	3
gb	4 2 1	2 3 4	3 1 2	4 1 2	1 4	3
g/	1 2 1	2 3 4	2 4 2	3 1 2	1 3	4

XI

	a b c	*d e f*	*g h i*	*k l m*	*n o*	*u*
	1 2 1	2 1 2	3 4 3			
23						
b/	1 3 1	2 1 2	3 4 3	identical with V		
bi	4 2 1	2 1 2	3 4 3	3 4 3	4 2	1
bg	4 2 1	2 1 2	3 1 3	3 4 3	4 2	1
bd	1 2 4	2 1 2	3 4 3			
14						
ec	1 2 4	3 1 2	3 4 3	1 2 3	1 2	4
ea	1 3 4	3 1 2	3 4 3	1 2 3	1 2	4
eh/	1 2 4	2 4 2	3 1 3	3 1 3	1 2	4
e/	1 2 4	2 4 2	3 4 3	3 1 3	1 2	4

(2, 4) 555555***



$b = d = 2,\ e = g = 3,\ a = 1,\ h = j.$

h = j	*c*	*f*	*i*	*k l m*	*n o p*	*u*	*L*
2	1	1	1	4 1 4	2 1 4	3	
2	1	1	3	4 1 4	2 1 4	3	
2	1	1	4	4 1 4	2 4 1	3	
2	1	2	1				*III*
2	1	2	3	4 1 4	1 4 1	3	
2	1	2	4	4 1 4	1 4 1	3	
2	1	4	1				*V*
2	1	4	3	4 1 2	1 4 1	3	

h=j	*c*	*f*	*i*	*k l m*	*n o p*	*u*	*V*
2	1	4	4	4 1 2	1 4 1	3	
2	3	1	3	4 1 4	2 4 1	3	
2	3	1	4	4 1 4	2 4 1	3	
2	3	2	3	4 1 4	1 4 1	3	
2	3	2	4	4 1 4	1 4 1	3	
2	3	4	3	4 1 2	1 4 1	3	
2	3	4	4	4 1 2	1 4 1	3	
2	4	1	4				*VII*
2	4	2	4				*VI*
2	4	4	4	1 4 2	1 4 1	3	
4	1	1	1	4 1 2	4 1 2	3	
4	1	1	2	4 1 2	4 2 1	3	
4	1	1	3	4 1 2	4 2 1	3	
4	1	2	1				*IV*
4	1	2	2	4 1 4	1 2 1	3	
4	1	2	3	4 1 4	1 2 1	3	
4	1	4	2	4 1 2	1 2 1	3	
4	1	4	3	4 1 2	1 2 1	3	
4	3	1	2	4 1 2	4 2 1	3	
4	3	1	3	4 1 2	4 2 1	3	
4	3	2	2	4 1 4	1 2 1	3	
4	3	2	3	4 1 4	1 2 1	3	
4	3	4	2	4 1 2	1 2 1	3	
4	4	1	2	1 4 2	4 2 1	3	
4	4	2	2	1 4 1	4 2 1	3	

I	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
	1 2 1	2 1 4 3	2 3 2			
23 *db*	1 2 4	2 1 4 3	2 3 2	1 4 2	1 4 1	3
dg	1 2 1	2 4 1 3	2 3 2	4 1 2	4 1 4	3
d/	1 2 1	3 1 4 3	2 3 2			
13 *eg*	1 2 1	3 1 2 3	2 3 2	4 2 4	1 4 1	3
ei	1 2 1	3 1 2 3	4 3 2	4 2 4	1 2 1	3
ea	1 4 1	3 1 4 3	2 3 2	2 4 2	1 4 1	3
e/	1 2 3	1 3 4 3	2 3 2	4 2 1	2 4 1	3

II	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
	1 2 1	2 3 2 1	2 3 2			
23 *e/*	1 2 1	3 2 3 1	2 3 2	4 1 4	2 4 1	3
eh	1 2 1	2 3 2 4	2 3 2			

II	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
^{12}*fh*	1 2 1	2 3 2 3	2 3 2	4 1 4	1 4 1	3
fd	1 2 1	2 4 2 4	2 3 2	3 1 3	1 3 1	4
f/	1 2 1	2 3 1 4	2 3 2	4 1 4	2 1 4	3
eb/	1 3 4	3 2 3 1	2 3 2			
^{13}*df*	1 3 4	3 4 3 1	2 3 2	1 2 1	2 3 1	4
di	1 3 4	3 4 3 1	4 3 2	1 2 1	2 3 1	4
db/	3 1 4	1 2 3 1	2 3 2	2 3 1	2 3 1	4
d/	1 3 4	1 2 3 1	2 3 2	2 3 1	2 3 1	4

III	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
	1 2 1	2 3 2 3	2 1 2			
^{13}*i/*	1 2 1	2 3 2 3	2 3 2	4 1 4	1 4 1	3
ic	1 4 1	2 3 2 3	2 1 4	3 1 4	1 4 3	2
ie	1 2 1	2 3 4 3	4 1 2	3 1 2	1 2 3	4
iag/	3 2 1	2 3 2 1	2 3 2	symmetrical to II		
ig/	1 2 1	2 3 2 1	2 3 2	identical with II		
ia/	3 2 1	2 3 2 3	2 3 2			
^{34}*ie*	3 2 1	2 3 1 3	1 3 2	3 4 2	4 2 4	1
ia/	4 2 1	2 3 2 3	2 4 2	4 1 4	1 4 1	3
ig,ea/	4 1 2	1 4 2 3	2 3 2	4 2 1	4 1 4	3
ig,e/	3 2 1	2 4 2 3	1 3 2			
^{23}*fd*	3 2 1	2 1 2 3	1 3 2	3 4 3	4 2 4	1
fi	3 2 1	2 4 2 3	4 3 2	3 1 3	1 2 1	4
f/	3 2 1	2 4 3 2	1 3 2	4 3 2	4 3 4	1
i/	3 2 1	2 3 2 3	2 4 2			
^{12}*bj*	4 2 1	2 3 2 3	2 4 2	4 1 4	1 4 1	3
bh	4 2 1	2 3 2 3	2 3 2	4 1 4	1 4 1	3
bf/	3 1 2	1 4 1 3	2 4 2	3 2 3	2 1 3	4
b/	3 1 2	1 3 2 3	2 4 2			
^{13}*ge*	3 1 2	1 3 4 3	2 4 2	3 2 1	2 1 3	4
ga	3 1 2	1 3 2 3	4 2 4	3 4 1	4 1 3	2
g/	3 1 2	1 3 2 1	2 4 2			
^{23}*hf*	3 1 2	1 3 2 4	2 4 2	3 2 1	3 1 3	4
hc	3 1 2	4 3 2 4	2 4 2	3 2 1	3 1 3	4
hj/	2 1 2	1 3 2 1	3 4 3	3 2 1	3 2 1	4
h/	3 1 2	1 3 2 1	3 4 2	3 2 1	3 2 1	4

IV		*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
		1 2 1	2 3 2 3	4 1 4			
	24 *b/*	1 4 1	2 3 2 3	4 1 4	3 4 1	4 1 3	2
	bh/	3 4 1	2 3 2 3	2 3 4	3 1 4	1 4 1	2
	bf/	1 4 1	4 3 4 3	4 1 4	identical with III		
	bfh/	1 4 1	2 3 4 1	2 1 4	3 4 1	3 4 3	2
	bd	1 2 3	2 3 2 3	4 1 4	1 4 1	4 1 2	3
	bj	3 2 1	2 3 2 3	4 1 4			
	34 *eg*	3 2 1	2 3 1 3	4 1 4	3 4 2	4 2 3	1
	ea	3 1 2	1 3 2 3	4 1 4	3 4 1	4 1 3	2
	e/	3 2 1	2 4 2 3	4 1 4			
	28 *fd*	3 2 1	2 1 2 3	4 1 4	3 4 3	4 2 3	1
	fa	3 2 1	2 4 2 3	1 4 1	3 1 3	1 2 3	4
	f/	3 2 1	2 4 3 2	4 1 4	4 3 2	4 3 2	1

V		*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
		1 2 1	2 3 4 3	2 1 2			
	18 *ce*	1 2 1	4 3 4 3	2 1 2	3 2 1	2 1 3	4
	cgi/	1 2 1	2 3 4 3	4 1 2	3 1 2	1 2 3	4
	cga/	1 2 1	2 3 4 3	4 3 4	4 1 2	1 2 1	3
	cg/	1 2 3	2 3 4 1	2 1 2	symmetrical to I		
	ci,eg	1 2 1	2 3 2 3	2 1 2	identical with III		
	ci,e/	1 4 1	2 1 4 3	2 1 4	3 4 3	1 4 3	2
	ca,ig	1 2 1	2 3 4 3	4 1 2	3 1 2	1 2 3	4
	ca,ie/	1 2 1	2 1 4 3	2 3 2	identical with I		
	ca,i/	1 2 1	2 3 4 3	2 3 2	4 1 2	1 4 1	3
	c/	1 2 3	2 3 4 3	2 1 2	1 4 1	2 1 4	3

VI		*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
		1 2 4	2 3 2 3	2 4 2			
	12 *hj*	1 2 4	2 3 2 3	2 3 2	1 4 1	4 1 4	3
	hd	1 2 4	2 4 2 4	2 4 2	3 1 3	1 3 1	4
	hf/	1 2 4	2 3 1 4	1 4 2	3 1 2	3 2 3	4
	h/	1 2 4	2 3 2 3	1 4 2			
	24 *f/*	1 2 4	2 3 4 3	1 4 2	3 1 2	1 2 3	4
	fi	1 2 4	2 3 2 1	3 4 2	1 4 1	4 2 1	3
	fd	1 2 4	2 1 2 3	1 4 2			

VI	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
^{23} *df*	1 2 4	2 4 2 3	1 4 2	3 1 3	1 2 3	4
dj	4 2 1	2 1 2 3	1 4 2	3 4 3	4 2 3	1
db/	1 3 4	3 1 2 3	1 4 2	1 2 3	1 2 3	4
d/	1 2 4	3 1 2 3	1 4 2	1 2 3	1 2 3	4

VII	*a b c*	*d e f g*	*h i j*	*k l m*	*n o p*	*u*
	1 2 4	2 3 1 3	2 4 2			
^{14} *fa*	1 2 4	2 3 1 2	3 4 3	3 1 2	3 1 2	4
fic/	1 2 1	2 3 4 3	2 1 2	identical with V		
fi/	1 2 4	2 3 4 3	2 1 2	1 4 1	2 1 4	3
fc/	1 2 1	2 3 4 3	2 4 2	4 1 2	1 4 1	3
f/	1 2 4	2 3 4 3	2 4 2	1 4 2	1 4 1	3

(3) Theorem: *In a minimal map the inequality*

$$3p_6 + p_7 \geq p_9 + \sum_{i>9} (i-9)p_i + 24$$

holds.

(3, 1) *Proof*: Let d denote the number of sides common to pentagons and polygons having more than five sides.

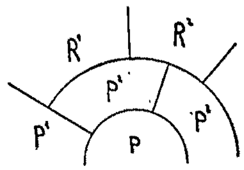
(3, 1, 1) By Birkhoff[6] there is not in a minimal map a pentagon touching three succeeding pentagons.

This implies that in a minimal map every pentagon touches at least two polygons having more than five sides; hence

(3, 1, 2) $$d \geq 2p_5.$$

(3, 2) By Franklin's theorem (1, 2, 2) in a minimal map no polygon having more than five sides touches pentagons only.

By a *row* we shall denote a configuration consisting of the polygon P—the *centre* of the row—and of the chain of polygons $P^1, P^2, \cdots, P^k$ $(k > 3)$ touching P where P^1 and P^k are not pentagons but P^i $(1 < i < k)$ are.



P^2 and P^{k-1}, the ends of the row, touch at least three polygons of more than five sides since R^1 and R^2 (see fig.) by (3, 1, 1) cannot be both pentagons. Every row has two ends, but a given pentagon can be an end in at most two rows. This is easy to be seen from the fig., in which P^2 can be the end (besides of the row with centre P) at most of one of the rows, with centres P^1, R^1 and R^2.

(3, 2, 1) Hence the number of rows cannot exceed the number of pentagons touching three or more polygons of more than five sides.

---

[6] *Loc. cit.* [2].

(3, 3) Let us denote by $\bar{q}_i$ the number of i-gons $\bar{Q}_i$ central in one row, by $\bar{\bar{q}}_i$ the number of polygons $\bar{\bar{Q}}_i$ central in at least two rows and by q_i the number of polygons Q_i which are not central in any row. Evidently for every $i > 5$ we have

(3, 3, 1) $$q_i + \bar{q}_i + \bar{\bar{q}}_i = p_i.$$

From (3, 1, 2) and (3, 2, 1) it follows that

$$d \geq 2p_5 + \sum_{i>5} \bar{q}_i + 2 \sum_{i>5} \bar{\bar{q}}_i$$

and by the well-known formula [7]

$$p_5 = \sum_{i>5} (i-6) p_i + 12,$$

we obtain

$$d \geq 2\Big[\sum_{i>6} (i-6) p_i + 12\Big] + \sum_{i>5} \bar{q}_i + 2 \sum_{i>5} \bar{\bar{q}}_i.$$

By (3, 3, 1) this becomes

(3, 3, 2) $$d \geq 2q_7 + 4q_8 + 6q_9 + \bar{q}_6 + 3\bar{q}_7 + 5\bar{q}_8 + 7\bar{q}_9 + 2\bar{\bar{q}}_6 + 4\bar{\bar{q}}_7 + 6\bar{\bar{q}}_8 + 8\bar{\bar{q}}_9 + \sum_{i>9} [(2i-12) q_i + (2i-11) \bar{q}_i + (2i-10) \bar{\bar{q}}_i] + 24.$$

(3, 4) On the other hand every Q_i $(i > 7)$ touches according to its definition at most $i - 4$ pentagons and every Q_6 and Q_7 touches at most 3 pentagons. According to (1, 2, 2) in every minimal map every $\bar{Q}_i$ $(i > 9)$ and $\bar{\bar{Q}}_i$ $(i \geq 6)$ touches at most $i - 2$ pentagons and every $\bar{Q}_i$ $(6 \leq i \leq 9)$ touches by theorem (2) at most $i - 3$ pentagons. We thus obtain

$$d \leq 3q_6 + 3q_7 + \sum_{i>7} (i-4) q_i + \sum_{i>9} (i-2)(\bar{q}_i + \bar{\bar{q}}_i) + \sum_{6 \leq i \leq 9} [(i-3)\bar{q}_i + (i-2)\bar{\bar{q}}_i].$$

(3, 5) According to (3, 3, 2) we have

$$3q_6 + 3q_7 + 4q_8 + 5q_9 + 3\bar{q}_6 + 4\bar{q}_7 + 5\bar{q}_8 + 6\bar{q}_9 + 4\bar{\bar{q}}_6 + 5\bar{\bar{q}}_7 + 6\bar{\bar{q}}_8 + 7\bar{\bar{q}}_9 + \sum_{i>9} [(i-4) q_i + (i-2)\bar{q}_i + (i-2)\bar{\bar{q}}_i] \geq 2q_7 + 4q_8 + 6q_9 + \bar{q}_6 + 3\bar{q}_7 + 5\bar{q}_8 + 7\bar{q}_9 + 2\bar{\bar{q}}_6 + 4\bar{\bar{q}}_7 + 6\bar{\bar{q}}_8 + 8\bar{\bar{q}}_9 + \sum_{i>9} [(2i-12) q_i + (2i-11)\bar{q}_i + (2i-10)\bar{\bar{q}}_i] + 24,$$

or

$$3q_6 + q_7 + 2\bar{q}_6 + \bar{q}_7 + 2\bar{\bar{q}}_6 + \bar{\bar{q}}_7 \geq q_9 + \bar{q}_9 + \bar{\bar{q}}_9 + \sum_{i>9} [(i-8) q_i + (i-9)\bar{q}_i + (i-8)\bar{\bar{q}}_i] + 24.$$

By (3, 3, 1) we finally obtain

$$3p_6 + p_7 \geq p_9 + \sum_{i>9} (i-9) p_i + 24.$$

EINSTEIN INSTITUTE OF MATHEMATICS,
HEBREW UNIVERSITY,
JERUSALEM, PALESTINE.

---

[7] E. g. Alfred Errera, *loc. cit.* [1], p. 29.

POLYNOMIALS WHOSE ROOTS LIE IN A SECTOR.*[1]

By LOUIS WEISNER.

1. Introduction.

From two polynomials

$$A(z) = a_0 + a_1 z + \cdots + a_n z^n, \qquad b_m \neq 0,$$
$$B(z) = b_0 + b_1 z + \cdots + b_m z^m,$$

we form the composition-polynomials

$$f(z) = a_0 b_0 + a_1 b_1 z + a_2 b_2 z^2 + \cdots + a_s b_s z^s,$$
$$g(z) = a_0 b_0 + 1!\, a_1 b_1 z + 2!\, a_2 b_2 z^2 + \cdots + s!\, a_s b_s z^s,$$

where $s = \min(m, n)$. The precise degree of $A(z)$ may be $< n$, in which case $A(z)$ has one or more infinite roots. Where it is necessary to take account of the zero or infinite roots of $A(z)$, or the zero roots of $B(z)$, we write

$$(1) \quad A(z) = a_p z^p + \cdots + a_q z^q, \qquad a_p a_q \neq 0;\ 0 \leq p \leq q \leq n,$$
$$(2) \quad B(z) = b_r z^r + \cdots + b_m z^m, \qquad b_r b_m \neq 0;\ 0 \leq r \leq m.$$

We shall assume that $f(z)$ and $g(z)$ do not vanish identically.

Malo has proved that if the roots of $A(z)$ are real, while those of $B(z)$ are real and of like sign, then the roots of $f(z)$ are real;[2] and Schur has proved that, under the same circumstances, the roots of $g(z)$ are real.[3]

I propose to generalize the theorems of Malo and Schur to the case in which the roots of $A(z)$ are real, while those of $B(z)$ lie in a sector with vertex at the origin and aperture $\leq \pi$. The term *sector* is employed in this restricted sense throughout this paper. If S is a sector, $-S$ denotes the sector consisting of the negatives of the numbers which belong to S, and $\pm S$ denotes the *double sector* consisting of S and $-S$.

We shall prove that if the roots of $A(z)$ are real and those of $B(z)$ lie in a sector S, then the roots of $f(z)$ and $g(z)$ lie in the double sector $\pm S$. When S consists of the positive or negative real axis, these results reduce to the theorems of Malo and Schur. Malo and Schur proved their respective

---

* Received March 13, 1941; Revised May 6, 1941.

[1] Presented to the American Mathematical Society, May 2, 1941.

[2] E. Malo, "Note sur les équations algébriques dont toutes les racines sont réeles," *Journal de Mathématiques spéciales* (4), vol. 4 (1895), p. 7; Pólya and Szegö, *Aufgaben und Lehrsätze aus der Analysis* II, p. 66, problem 155.

[3] J. Schur, "Zwei Sätze über algebraische Gleichungen mit lauter reelen Wurzeln," *Journal für Mathematik*, vol. 144 (1914), 75-88; Pólya and Szegö, *loc. cit.*, problem 156.

theorems by setting up suitable Sturm sequences. As Sturm's Theorem is inapplicable to the imaginary roots of a polynomial, it was necessary to resort to other methods to generalize their results. Appropriate tools were found in the Gauss-Lucas Theorem and in a theorem of Laguerre. The proofs of the generalized theorems are simpler than the original proofs of Malo and Schur although the same formulas are employed up to a certain point.

Professor G. Pólya, who read the manuscript of this paper, pointed out to me the possibility of establishing a much sharper composition-theorem than the two mentioned, from which these theorems may be derived. At his suggestion I have retained my own proofs and have added a section including his contribution.

2. Generalization of Malo's Theorem.

LEMMA 1.[4] *Let* $F(z) = c_0 + c_1 z + \cdots + c_t z^t \not\equiv 0$ *be a polynomial whose roots lie in a half-plane whose boundary passes through the origin, and suppose that not all the roots of* $F(z)$ *lie on the boundary of the half-plane. If* c_u *is the first and* c_v *the last non-vanishing coefficient of* $F(z)$*, then* $c_k \neq 0$, $u \leq k \leq v$.

If $u = v$, $F(z) = c_v z^v$, all of whose roots lie on the boundary L[5] of the half-plane in question; hence $u < v$. By the Gauss-Lucas Theorem, a finite root of $F^{(k)}(z)$ which lies on L must be a $(k+1)$-fold root of $F(z)$. Since 0 is a u-fold root of $F(z)$,

$$F^{(k)}(0) \neq 0, \qquad u \leq k \leq v.$$

In what follows we use the notation and terminology of § 1.

THEOREM 1. *If the roots of* $A(z)$ *are real, and the roots of* $B(z)$ *lie in a sector* S*, and if* $f(z) \not\equiv 0$*, then the roots of* $f(z)$ *lie in the double sector* $\pm S$.[6]

Since $f(z) \not\equiv 0$, $q \geq r$ and $m \geq p$. If $q = r$ or $m = p$, $f(z)$ consists of just one non-vanishing term, and the theorem is trivial. Now let $q > r$ and $m > p$. Suppose that $f(z)$ has a root ζ outside $\pm S$. Since the roots of $A(z)$ are real, the roots of the polynomial

$$z^n A(\zeta z^{-1}) = a_n \zeta^n + a_{n-1}\zeta^{n-1} z + \cdots + a_1 \zeta z^{n-1} + a_0 z^n$$

lie on the line through 0 and ζ. One of the half-planes of which this line is the boundary includes the sector S. Therefore all the roots of the polynomial $F(z) = z^n A(\zeta z^{-1}) B(z)$

---

[4] A different proof is given by E. Laguerre, *Oeuvres*, vol. 1, pp. 135-136.

[5] An infinite root lies on every line, in every half-plane, and in every sector.

[6] The theorem is trivial when S is a half-plane.

$$= a_n b_0 \zeta^n + \cdots + (a_0 b_0 + a_1 b_1 \zeta + \cdots + a_s b_s \zeta^s) z^n + \cdots + a_0 b_m z^{n+m}$$

lie in this half-plane, and not all the roots lie on its boundary. Since the coefficient of z^n in $F(z)$ is $f(\zeta) = 0$, two possibilities arise according to the lemma: either all preceding coefficients vanish, or all succeeding coefficients vanish. Neither of these alternatives is realized since the coefficient of z^{n-q+r} is $a_q b_r \zeta^q \neq 0$, and the coefficient of z^{n-p+m} is $a_p b_m \zeta^{n-p} \neq 0$. We conclude that ζ lies in $\pm S$.

Corollary.[7] *If the roots of $A(z)$ are real and negative, a sector which includes the roots of $B(z)$ includes the roots of $f(z)$.*[8]

The roots of $z^n A(\zeta z^{-1})$ now lie on the ray from the origin which passes through $-\zeta$. If ζ, a root of $f(z)$, lies in $-S$, the ray in question is included in S. Therefore the roots of the polynomial $F(z)$ lie in S. Since the coefficient of z^n in $F(z)$ vanishes, a contradiction arises as in the proof of Theorem 1. Therefore ζ lies in S.

3. Generalization of Schur's Theorem. From the polynomials $A(z)$ and $B(z)$ we form the polynomial

$$\phi(z) = a_n B(z) + a_{n-1} B'(z) + \cdots + a_0 B^{(n)}(z) \not\equiv 0. \tag{3}$$

It is known that *a half-plane whose boundary passes through the origin, and which includes the roots of $A(z)$ and $B(z)$, includes the roots of $\phi(z)$.*[9] We proceed to investigate the circumstances under which $\phi(z)$ has a root on the boundary of the half-plane.

With the notation (1) for $A(z)$, set $C(z) = B^{(n-q)}(z)$, so that

---

[7] A direct proof of the corollary may be given by means of a theorem of I. Schur cited by G. Szegö, "Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln Algebraischer Gleichungen," *Mathematische Zeitschrift*, vol. 13 (1922), p. 37. See also Pólya and Szegö, *loc. cit.*, p. 65, problem 153.

[8] A special case is established by A. Cohn, "Über die Anzahl der Wurzeln einer algebraischer Gleichung in einem Kreise," *Mathematische Zeitschrift*, vol. 14 (1922), p. 148.

[9] See Eduard Benz, "Über lineare, verschiebungstreue Funktionaloperationen und die Nullstellen ganzer Funktionen," *Commentarii Mathematici Helvetici*, vol. 7 (1935), p. 248. General theorems concerning the roots of $\phi(z)$ where circular regions are emphasized are given by T. Takagi, "Note on the algebraic equations," *Proceedings of the Physico-Mathematical Society of Japan*, vol. 3 (1921), 175-179; J. L. Walsh, "On the location of the roots of certain types of polynomials," *Transactions of the American Mathematical Society*, vol. 24 (1922), p. 173; J. L. Walsh, "On the location of the roots of polynomials," *Bulletin of the American Mathematical Society*, vol. 30 (1924), p. 51; L. Berwald, "Über die Lage der Nullstellen von Linearkombinationen eines Polynoms und seiner Ableitungen in Bezug auf einen Punkt," *Tôhoku Mathematical Journal*, vol. 37 (1933), p. 67.

$$\phi(z) = a_q C(z) + a_{q-1} C'(z) + \cdots + a_p C^{(q-p)}(z).$$

The precise degree of $\phi(z)$ is $t = m - n + q$ which we take ≥ 1. Let α be a finite root of $A(z)$ and ζ a root of the polynomial

$$\psi(z) = C(z) - \alpha C'(z).$$

The centroid of $C(z)$ relative to ζ is

$$\zeta - tC(\zeta)/C'(\zeta) = \zeta - t\alpha.$$

If ζ lies on the boundary L of the half-plane K in question, but α does not, then $\zeta - t\alpha$ is outside K;[10] consequently all the roots of $C(z)$ are equal and lie on L.[11] Writing

$$C(z) = c(z - \gamma)^t, \qquad c \neq 0,$$

we have

$$\psi(z) = c(z - \gamma)^{t-1}(z - \gamma - t\alpha).$$

Therefore $\zeta = \gamma$ or $\zeta = \gamma + t\alpha$. In the first case ζ is a multiple root of $C(z)$; in the second case α lies on L since ζ and γ lie on L.

Thus if ζ is on L and is not a multiple root of $C(z)$, α is also on L; hence $\zeta - t\alpha$ is on L. If $\alpha = 0$, $\psi(z) = C(z)$, and ζ is a simple root of $C(z)$. If $\alpha \neq 0$, ζ and $\zeta - t\alpha$ are distinct, and every circle (in particular the line L) through ζ and $\zeta - t\alpha$ separates the roots of $C(z)$. Since all the roots of $C(z)$ lie in K, they must lie on L. Conversely, if α and all the roots of $C(z)$ lie on L, so do the roots of $\psi(z)$. We conclude that a finite root ζ of $\psi(z)$ lies on L if and only if ζ is a multiple root of $C(z)$, or if $\alpha = 0$ and ζ is a simple root of $C(z)$ which lies on L, or if α and all the roots of $C(z)$ lie on L.

Now if $\alpha_1, \cdots, \alpha_q$ are the finite roots of $A(z)$, then

$$\phi(z) = a_q \left(1 - \alpha_1 \frac{d}{dz}\right) \cdots \left(1 - \alpha_q \frac{d}{dz}\right) C(z).$$

Therefore by repeated applications of the preceding result we have the following

LEMMA 2. *Let K be a half-plane whose boundary L passes through the origin and which includes all the roots of the polynomials $A(z)$ and $B(z)$. A finite root of the polynomial $\phi(z)$ defined by* (3) *lies on L if and only if it is a root of $C(z) = B^{(n-q)}(z)$ and the roots of $A(z)$ are* 0 *or* ∞; *or if it is a root of $C(z)$ of multiplicity $\geq n - p + 1$; or if all the roots of $A(z)$ and $B(z)$ lie on L, in which case all the roots of $\phi(z)$ lie on L.*[12]

---

[10] E. Laguerre, *Oeuvres*, vol. 1, pp. 57, 134.

[11] Pólya and Szegö, *loc. cit.*, p. 57, problem 107.

[12] The last case includes a well-known theorem of Hermite, according to which the roots of $\phi(z)$ are real if those of $A(z)$ and $B(z)$ are real.

THEOREM 2. *If the roots of $A(z)$ are real, and the roots of $B(z)$ lie in a sector S, and if $g(z) \not\equiv 0$, then the roots of $g(z)$ lie in the double sector $\pm S$.*

Suppose $g(z)$ has a root ζ outside $\pm S$. As in the proof of Theorem 1, the roots of $z^n A(\zeta z^{-1})$ and $B(z)$ lie in a half-plane K whose boundary L passes through the origin. Moreover, the finite roots of $B(z)$ lie on L if and only if they are all equal to 0. Now K also includes the roots of the polynomial

$$\phi(z) = a_0 B(z) + a_1 \zeta B'(z) + \cdots + a_n \zeta^n B^{(n)}(z),$$

and

$$\phi(0) = a_0 b_0 + 1!\, a_1 b_1 \zeta + \cdots + s!\, a_s b_s \zeta^s = g(\zeta) = 0.$$

Since $\phi(z)$ has a root 0 on L, we may apply Lemma 2 observing that, since

$$z^n A(\zeta z^{-1}) = a_q \zeta^q z^{n-q} + \cdots + a_p \zeta^p z^{n-p},$$

p and q are to be replaced by $n - q$ and $n - p$ respectively. The following cases arise:

(1). 0 is a root of $B^{(p)}(z)$, and the roots of $z^n A(\zeta z^{-1})$ are 0 or ∞; hence $p = q$ and $A(z) = a_p z^p$. Therefore

$$g(z) = p!\, a_p b_p z^p,$$

and the supposition that ζ is outside $\pm S$ is false.

(2). 0 is a root of $B^{(p)}(z)$ of multiplicity $\geq q + 1$. Here $g(z) \equiv 0$.

(3). All the roots of $B(z)$ equal 0 or ∞. Here $g(z)$ consists of only one non-vanishing term, and the supposition that ζ is outside $\pm S$ is false.

We conclude that ζ lies in $\pm S$.

COROLLARY.[13] *If the roots of $A(z)$ are real and negative, a sector which includes the roots of $B(z)$ includes the roots of $g(z)$.*

The proof is similar to that of the corollary to Theorem 1.

4. Extension of Theorems 1 and 2.[14] We now write

$$A(z) = a_0 + \binom{n}{1} a_1 z + \binom{n}{2} a_2 z^2 + \cdots + a_n z^n,$$

$$B(z) = b_0 + \binom{n}{1} b_1 z + \binom{n}{2} b_2 z^2 + \cdots + b_n z^n,$$

and from them form the composition-polynomial

$$h(z) = a_0 b_0 + \binom{n}{1} a_1 b_1 z + \binom{n}{2} a_2 b_2 z^2 + \cdots + a_n b_n z^n.$$

---

[13] See footnote 7.

[14] As noted in § 1, the results of this section are due to G. Pólya.

THEOREM 3. *If the roots of $A(z)$ are real, and the roots of $B(z)$ lie in a sector S, and if $h(z) \not\equiv 0$, then the roots of $h(z)$ lie in the double sector $\pm S$.*

Suppose $h(z)$ has a root ζ not in $\pm S$. As in the proof of Theorem 1, the roots of $z^nA(-\zeta z^{-1})$ and $B(z)$ lie in a half-plane K whose boundary L passes through the origin. The theorem being trivial when $B(z)$ consists of just one non-vanishing term, we shall suppose that at least one root of $B(z)$ is interior to K. Now the equation $h(\zeta) = 0$ asserts that the polynomials $z^nA(\zeta z^{-1})$ and $B(z)$ are apolar. Therefore, by Grace's Apolarity Theorem,[15] at least one root of $z^nA(-\zeta z^{-1})$ is outside K. Since the roots of this polynomial lie on L, the supposition that ζ is outside $\pm S$ is false.

Theorems 1 and 2 may be derived from Theorem 3 by means of the following theorem due to Laguerre:[16] *If all the roots of the polynomial*

$$c_0 + c_1z + c_2z^2 + \cdots + c_nz^n$$

are real, the same is true of the roots of the polynomial

$$\frac{c_0}{n!} + \frac{c_1z}{(n-1)!} + \frac{c_2z^2}{(n-2)!} + \cdots + c_nz^n.$$

Since the roots of $A(z)$ are real, so are the roots of the polynomial

$$\frac{a_0}{n!} + \frac{\binom{n}{1}a_1z}{(n-1)!} + \frac{\binom{n}{2}a_2z^2}{(n-2)!} + \cdots + a_nz^n.$$

Applying Theorem 3 to this polynomial and $B(z)$, Theorem 2 results. By applying the cited theorem of Laguerre twice it follows that the roots of the polynomial

$$\frac{a_0}{n!} + \frac{\binom{n}{1}a_1z}{1!(n-1)!} + \frac{\binom{n}{2}a_2z^2}{2!(n-2)!} + \cdots + \frac{a_nz^n}{n!}$$

are real. Applying Theorem 3 to this polynomial and $B(z)$, Theorem 1 results.

HUNTER COLLEGE OF THE CITY OF NEW YORK.

---

[15] J. H. Grace, "The zeros of a polynomial," *Proceedings of the Cambridge Philosophical Society*, vol. 11 (1902), 352-357; Pólya and Szegö, *loc. cit.*, p. 64, problem 145. The name *Grace's Apolarity Theorem* was first proposed by E. B. Van Vleck, "On the location of roots of polynomials and entire functions," *Bulletin of the American Mathematical Society*, vol. 35 (1929), p. 650.

[16] E. Laguerre, *Oeuvres* I, p. 31; Pólya and Szegö, *loc. cit.*, p. 47, problem 65.

SOLUTION OF A PROBLEM OF AYRES.*

By Loyal F. Ollmann

In a paper in 1932, W. L. Ayres[1] suggests the question: If K is any finite subset of a Peano space M which is a subset of the Euclidean plane, and if no two points of K may be separated[2] by the omission of any two points of M, then does there exist a simple closed curve J such that $M \supset J \supset K$? Ayres answers the question in the affirmative in the same paper for $K = 3$ points, and in later unpublished work he discovered an example which answers the question in the negative for $K \geqq 11$ points.

In this paper we prove that the answer to Ayres' question is in the affirmative for $K \leqq 5$ points and give examples to show that it is in the negative for $K > 5$ points. We also show that the same results are obtained provided the condition "no two points of K may be separated" is changed to read "no two points of M may be separated." Examples are constructed to show that in certain Peano spaces a finite set K can not be joined by m (where m is any given integer) mutually exclusive simple closed curves.

The writer wishes to express his thanks to Professor W. L. Ayres for his helpful criticisms and generous assistance in the preparation of this paper.

Historical note. The theorem, that if two points of a planar Peano space are not separated by the omission of any single point, then they can be joined by a simple closed curve, was proved by G. T. Whyburn[3] and W. L. Ayres.[4] K. Menger[5] proved the theorem for regular curves while Ayres[6] extended the

---

* Received January 27, 1941. Presented to the American Mathematical Society, April 15, 1939. This paper is a portion of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Michigan.

[1] "On joining finite subsets of a Peano space by arcs and simple closed curves," *Fundamenta Mathematicae*, vol. 19 (1932), pp. 79-91.

[2] A point set H of a connected set M is said to *separate* the points x and y in M if $M - H$ is the sum of two mutually separated sets containing x and y respectively.

[3] "Some properties of continuous curves," *Bulletin of the American Mathematical Society*, vol. 33 (1927), pp. 305-308.

[4] This result was announced in an abstract "On the separation of points of a continuous curve by arcs and simple closed curves," *Bulletin of the American Mathematical Society*, vol. 33 (1927), p. 266.

[5] "Zur allegemeinen Kurventheorie," *Fundamenta Mathematicae*, vol. 10 (1927), pp. 96-115.

[6] "Concerning continuous curves in metric space," *American Journal of Mathematics*, vol. 51 (1929), pp. 577-594.

result to general Peano spaces. Conditions under which three points lie on a simple closed curve have been given by Ayres.[7]

The theorem, that if two closed subsets A and B of a Peano space are not separated by any n points, then they can be joined by $n+1$ independent arcs,[8] was proved by N. E. Rutt[9] for the case where M is a subset of the plane and A and B are single points. Menger[10] proved the theorem for regular curves. It was extended to compact Peano spaces by G. Nöbeling[11] and to locally compact Peano spaces by Leo Zippin.[12] Ayres[13] proved that if K is a set consisting of n points $(n > 1)$ of a Peano space P such that no two points of K may be separated by the omission of $n-1$ points of P, then there exists a simple closed curve of P containing K.

Notation and preliminary lemmas. Throughout this paper we assume that K is a finite subset of a Peano space M, a subset of the Euclidean plane E_2. We say that a set K *is 3-connected in* M if there do not exist two points x and y of M such that $M-(x+y)=M_1+M_2$, where M_1 and M_2 are mutually separated and $M_1K \neq 0$, $M_2K \neq 0$. This term seems appropriate since between any two points of K there are three independent arcs in M. It is evident that if K is 3-connected in M, then every point of K is of Menger-Urysohn order $\geqq 3$. We also say that if for every two points x and y, the set $M-(x+y)$ is connected, then M *is 3-connected,* in which case every point of M is of order $\geqq 3$. The letter J will be used to denote a simple closed curve contained in M. In general, capital letters will denote sets of points, while lower case letters denote individual points or numbers.

LEMMA 1. *If K is 3-connected in M, and $p_1 \epsilon K$ is interior to the simple closed curve $J \subset M$, and $p_2 \epsilon K$ is exterior to J, then there exist in M three arcs ρ_1, ρ_2, and ρ_3 which join p_1 to J such that $\rho_i \cdot \rho_j$ is the point p_1 $(i \neq j)$.*

The lemma follows as a direct result of the theorem due to Rutt.[14] For let α_1, α_2, and α_3 be three independent arcs which join p_1 and p_2. By the Jordan curve theorem[15] each of these arcs has at least one point in common with J. Let a_i $(i = 1, 2, 3)$ be the first point in the order from p_1 to p_2 which α_i has in common with J. Since the α_i are independent, the points a_i are distinct. Let ρ_i be the subarc p_1a_i of α_i, then $\rho_i \cdot \rho_j = p_1$, and the lemma is proved.

---

[7] "Continuous curves which are cyclicly connected," *Bulletin de l'Académie Polonaise des Sciences et des Lettres*, (1928), pp. 127-142.

[8] Arcs which are mutually exclusive except possibly for their common end-points are called *independent* arcs.

[9] "Concerning the cut points of a continuous curve when the arc curve, AB, contains exactly N independent arcs," *American Journal of Mathematics*, vol. 51, No. 2 (1929), pp. 217-246.

[10] *Loc. cit.*

[11] "Eine Verschärfung des n-Beinsatzes," *Fundamenta Mathematicae*, vol. 18 (1932), pp. 23-38.

[12] "Independent arcs of a continuous curve," *Annals of Mathematics*, vol. 34, No. 1 (1933), pp. 95-113.

[13] "On joining finite subsets of a Peano space by arcs and simple closed curves," *loc. cit.*[6]

[14] *Loc. cit.*[9] Theorem IV.

LEMMA 2. *If K is 3-connected in M, and $p_1 \epsilon K$ is interior to the simple closed curve $J \subset M$, and p_2, $p_3 \epsilon K$ lie on J, then either there exist* (1) *three arcs ρ_1, ρ_2, and ρ_3 which join p_1 to J such that $\rho_i \cdot \rho_j = p_1$ $(i, j = 1, 2, 3; i \neq j)$, or* (2) *two arcs ρ_1 and ρ_2 such that $\rho_1 \cdot J = p_2$, $\rho_2 \cdot J = p_3$, and $\rho_1 \cdot \rho_2 = p_1$.*

Proof. Let α_i $(i = 1, 2, 3)$ denote three independent arcs from p_1 to p_2, and let a_i be the first point of J on α_i in the order p_1 to p_2. If the a_i are distinct points, the lemma is true for this case. If $a_i = a_j$ $(i \neq j)$, then it is clear from the definition of independent arcs that $a_i = a_j = p_2$. Let β_i $(i = 1, 2, 3)$ denote three independent arcs joining p_1 to p_3, with b_i the first point of J on β_i in the order p_1 to p_3. If the b_i are distinct points then the result is obvious. Therefore we may assume that $a_1 = a_2 = p_2$ and $b_1 = b_2 = p_3$. In the order p_3 to p_1 let b_1' be the first point of $\alpha_1 + \alpha_2$ on β_1, and suppose that b_1' belongs to α_1. Let $\rho_1 = \alpha_2$ and $\rho_2 =$ subarc p_1b_1' of $\alpha_1 +$ subarc $b_1'p_3$ of β_1, which satisfies conclusion (2) of the lemma.

LEMMA 3. *If K is 3-connected in M, and $p_1 \epsilon K$ is interior to the simple closed curve $J \subset M$, and p_2, p_3, and $p_4 \epsilon K$ lie on J, then there exist three arcs ρ_1, ρ_2, and ρ_3 which join p_1 to J such that $\rho_i \cdot \rho_j = p_1$ $(i \neq j)$.*

Proof. Using the theorem of Rutt, we let α_i $(i = 1, 2, 3)$ denote the three independent arcs from p_1 to p_2. Let a_i $(i = 1, 2, 3)$ represent the first points of α_i on J in the order p_1 to p_2. If the a_i are distinct points, then the conclusion of the lemma follows. If two of the points a_i coincide, then the independence of the arcs α_i requires the point of coincidence to be p_2. We may thus suppose $a_1 = a_2 = p_2$. It is not important in this discussion whether a_3 and p_2 coincide or not. Because of Lemma 2 there are either three independent arcs from p_3 and p_4 to p_1 (in which case the conclusion of this lemma is immediate) or there exist two independent arcs from p_3 and p_4 respectively to p_1. We assume that β joins p_3 to p_1 and γ joins p_4 to p_1. In the order p_3 to p_1 on β, let b represent the first point of β on $\Sigma\alpha_i$, and let c be the first point of γ on $\Sigma\alpha_i$ in the order p_4 to p_1. If b and c lie on different arcs, say α_1 and α_3, respectively, then p_1b is a subset of α_1, $bp_3 \subset \beta$, $p_1c \subset \alpha_3$, and $cp_4 \subset \gamma$. Thus $p_1b + bp_3$, $p_1c + cp_4$, and α_2 are the required arcs, and they would also satisfy the lemma if $b = p_1$ or $c = p_1$ or both $b, c = p_1$.

---

[15] C. Jordan, *Cours d'Analyse*, Paris (1893), p. 90.

We now assume that b and c lie on the same arc and for our purpose we assume this arc to be α_3. We further assume the order on α_3 to be p_1cba_3. In the order b to p_1 on the subarc bp_1 of β, let b_1 represent the first point on β of the point-set $\alpha_1 + \alpha_2$. In like manner let c_1 represent the first point of $\alpha_1 + \alpha_2$ on the arc cp_1 of γ in the order c to p_1. Either b_1 or c_1 or both may be the point p_1 but in these cases the arcs are quite evident. For example, if c_1 lies on α_2 and $b_1 = p_1$, then α_1, subarc p_1c_1 of α_2 + subarc c_1p_4 of γ, and β are the three desired arcs. The two cases remaining to be discussed are those in which b_1 and c_1 lie on different arcs or on the same arc.

We suppose that b_1 lies on α_2 and c_1 lies on α_1, both points being distinct from p_1. In the order p_1 to a_3 on α_3, let x denote the first point of α_3 on the sum of the subarcs cc_1 of γ and bb_1 of β. If x belongs to β then subarc p_1c_1 of α_1 + subarc c_1p_4 of γ, α_2, and subarc p_1x of α_3 + subarc xp_3 of β, are three independent arcs from p_1 to J. If x is on γ then subarc p_1x of α_3 + subarc xp_4 of γ, subarc p_1b_1 of α_2 + subarc b_1p_3 of β, and α_1 are the three desired arcs.

We now suppose that b_1 and c_1 both lie on α_2 (the case for α_1 is similar). Assume the order on α_2 to be $p_2b_1c_1p_1$. In the order p_1 to a_3 on α_3, let x denote the first point on p_1c of subarc c_1c of γ. The arc p_1x does not have a point in common with the subarc b_1p_3 as may be seen by application of the Jordan curve theorem to the various cases which may arise. The three arcs satisfying this case are: α_1, subarc p_1x of α_3 + subarc xp_4 of γ, and subarc p_1b_1 of α_2 + subarc b_1p_3 of β. If the order on α_2 is $p_2c_1b_1p_1$ let y be the first point in the order p_1 to a_3 on α_3 of subarc b_1b of β. As above, p_1y does not have a point in common with c_1c of γ. In this case the three arcs α_1, subarc p_1y of α_3 + subarc yp_3 of β, and subarc p_1c_1 of α_2 + subarc c_1p_4 of γ are the required arcs. This completes the proof of the lemma.

LEMMA 4. *If K is 3-connected in M, then for any three points p_1, p_2, $p_3 \epsilon K$, there exist three independent arcs α_1, α_2, and α_3 with end-points p_1 and p_2 such that p_3 belongs to one of the arcs α_i.*

Proof. Let α_1, α_2, and α_3 be any three independent arcs which join p_1 and p_2. If p_3 belongs to one of the arcs α_i, then the lemma is proved. Suppose then that p_3 is interior to one of the domains complementary to $\alpha_1 + \alpha_2 + \alpha_3$ whose boundary we assume to be $J' = \alpha_1 + \alpha_2$, a simple closed curve. By Lemma 2, there are either (1) three independent arcs β_1, β_2, and β_3 that join p_3 to J' such that the end-points b_i are distinct, or (2) two independent arcs δ_1 and δ_2 such that $\delta_1 \cdot J' = p_1$, $\delta_2 \cdot J' = p_2$, and $\delta_1 \cdot \delta_2 = p_3$. In (1) at least two of the points b_i belong to either α_1 or α_2. Suppose b_1 and

b_2 belong to α_1 in the order $p_1 b_1 b_2 p_2$. The point set $p_1 b_1 + \beta_1 + \beta_2 + b_2 p_2$ is an arc which contains p_3, and with α_2 and α_3 forms a set of independent arcs joining p_1 and p_2. In (2), we merely let $\delta_1 + \delta_2$ replace α_1 in the independent set α_i, which proves the lemma.

THEOREM 1. *Let K be any set of $\leqq 5$ points which is 3-connected in M, then there exists a simple closed curve J such that $M \supset J \supset K$.*

Proof. Let K be the five points $(p_1, p_2, \cdots, p_5)$. By Lemma 4 there exist three independent arcs, α_1, α_2, and α_3, in M which join p_1 and p_2 and contain p_3. We let α_1 be the arc which contains p_3. Let D_1, D_2, and D_3 be the complementary domains of $\alpha_1 + \alpha_2 + \alpha_3$ such that the frontier of D_i does not contain α_i. Because of the symmetry of the points p_i and the arcs α_i, it is necessary to discuss only the following cases:

Case *I*. The points p_4 and p_5 lie on the arcs α_i.
  I$_a$. One of the arcs α_2 or α_3 contains neither the point p_4 nor the point p_5.
  I$_b$. Each arc α_i contains one of the points.

Case *II*. One arc contains p_4, while p_5 is an interior point of one of the complementary domains D_i.
  II$_a$. p_4 is on α_1, p_5 is in D_1.
  II$_b$. p_4 is on α_1, p_5 is in D_2 or D_3.
  II$_c$. p_4 is on α_2, p_5 is in D_3.
  II$_d$. p_4 is on α_2, p_5 is in D_1 or D_2.

Case *III*. Both points are interior points of the complementary domains D_i.
  III$_a$. Both p_4 and p_5 are in D_2 or D_3.
  III$_b$. Both p_4 and p_5 are in D_1.
  III$_c$. p_4 is in D_3, p_5 is in D_2.
  III$_d$. p_4 is in D_2 or D_3, p_5 is in D_1.

Case I$_a$. Let α_3 be that arc which does not contain the points p_4 or p_5. Then $\alpha_1 + \alpha_2$ is a simple closed curve containing K.

Case I$_b$. Suppose α_2 contains p_4 and α_3 contains p_5. Let α'_1 and α''_1 denote the subarcs $p_3 p_1$ and $p_3 p_2$ of α_1. In the same manner, let p_4 divide α_2 into subarcs α'_2 and α''_2, and p_5 divide α_3 into α'_3 and α''_3. Due to symmetry we also suppose that D_2 is the unbounded domain. Since $p_4 \epsilon K$ is interior to the simple closed curve $\alpha_1 + \alpha_3$, there exist (by Lemma 3) three independent arcs β_1, β_2, and β_3 which join p_4 to distinct points b_1, b_2, and b_3, respectively, of $\alpha_1 + \alpha_3$. If two of the b_i belong to the same subarc α'_1, α''_1, α'_3, or α''_3, then it is evident that K is contained in a simple closed curve.

Suppose, however, that b_1 is on α'_1, b_2 is on α''_1, and b_3 is on α'_3. Every other distribution will be similar to this case. In the order from p_1 to p_4, let x' be the first point on α'_2 which is also a point of $\Sigma\beta_i$. Denote the subarc $x'p_4$ of β_i by β'_i. Then the following cases represent the possible distributions of x' on the β_i with a simple closed curve containing K, thus satisfying the theorem:

1) x' on β_1: $p_1x' + \beta'_1 + \beta_3 + b_3p_5 + \alpha''_3 + \alpha_1 = J$

2) x' on β_2: $p_1x' + \beta'_2 + \beta_1 + b_1p_3 + \alpha''_1 + \alpha_3 = J$

3) x' on β_3: $p_1x' + \beta'_3 + \beta_1 + b_1p_3 + \alpha''_1 + \alpha_3 = J$

This completes the proof for case I_b.

Case II_a. By Lemma 1, there exist three independent arcs β_1, β_2, and β_3 which join p_5 to the distinct points b_1, b_2, and b_3 of $\alpha_2 + \alpha_3$. Two of the points b_i must lie on α_2 or on α_3. Suppose b_1 and b_2 lie on α_2 in the order $p_1b_1b_2p_2$. Then let J *be* $\alpha_1 + p_1b_1 + \beta_1 + \beta_2 + b_2p_2$. The case for b_1 and b_2 on α_3 is similar.

Case II_b. Suppose p_5 is in D_3, the other case being similar to it. We also assume the order $p_1p_3p_4p_2$ on α_1. Let β_1, β_2, and β_3 be three independent arcs (Lemma 3) with end-points p_5 and b_i $(i = 1, 2, 3)$, where the b_i are dictinct points of the simple closed curve $\alpha_1 + \alpha_2$. If two of the b_i, say b_1 and b_2, lie on α_2 in the order $p_1b_1b_2p_2$, then $p_1b_1 + \beta_1 + \beta_2 + b_2p_2 + \alpha_1 = J$ is a simple closed curve containing K. Let α'_1, α''_1, and α'''_1 denote the subarcs p_1p_3, p_3p_4, and p_4p_2 of α_1. If b_1 and b_2 lie on α'_1 then $p_1b_1 + \beta_1 + \beta_2 + b_2p_3 + \alpha''_1 + \alpha'''_1 + \alpha_2$ is the desired simple closed curve J; and similarly for two of the b_i on the arcs α''_1 or α'''_1. We now assume that each of the arcs α'_1, α''_1, α'''_1, and α_2 contains at most one of the points b_i. If one end-point, say b_3, is on the arc α_2, then either b_1 or b_2 belongs to α'_1 or α'''_1 for otherwise both would belong to α''_1. Suppose that b_1 is on α'_1. Then K would be contained in the simple closed curve $p_1b_3 + \beta_3 + \beta_1 + b_1p_3 + \alpha''_1 + \alpha'''_1 + \alpha_3$. For b_2 on α'''_1, J is obtained similarly.

The case remaining is that of b_1 on α'_1, b_2 on α''_1, and b_3 on α'''_1. For the discussion of this case, assume that D_1 is the unbounded complementary domain of $\alpha_1 + \alpha_2 + \alpha_3$ for, if necessary, we can perform an inversion. By Lemma 3, let γ_1, γ_2, and γ_3 be three independent arcs with end-points p_4 and c_i $(i = 1, 2, 3)$, respectively, such that the c_i are distinct points on the simple closed curve $J' = \alpha_3 + \alpha'_1 + p_3b_2 + \beta_2 + \beta_3 + b_3p_2$ of which p_4 is an interior point. If two of the points c_i lie on the same subarc α_3, α'_1, p_3p_5, or p_5p_2 of J', then the simple closed curve which contains K is evident. If one of the points c_i, say c_1, is on $\langle\alpha_3\rangle$[16], then let x be the first point (possibly p_4), in the order from c_1 to p_4, which γ_1 has in common with the subarc b_2b_3 of α_1. Suppose x belongs to b_2p_4, then $\alpha'_1 + p_3b_2 + \beta_2 + \beta_3 + b_3p_4 + p_4x + xc_1 + c_1p_2 + \alpha_2$ is a simple closed curve containing K; and similarly for x belonging to p_4b_3. Thus when one of the c_i is on the arc $\langle\alpha_3\rangle$ there is a simple closed curve J which has properties required for the conclusion of the theorem.

We assume next that each of the subarcs p_1p_3, p_3p_5, and p_5p_2 of J' contains one of the c_i. Suppose c_1 is on p_1p_3, c_2 is on p_3p_5 and c_3 is on p_5p_2. For convenience in notation, and because the arcs b_2p_4 and b_3p_4 of α_1 do not enter our discussion, we let $c_2 = b_2$ and $c_3 = b_3$. We assume also that $c_1 = b_1$, which assumption merely narrows the availability of the points b_1 and c_1 as interior points of distinct arcs in a simple closed curve. Denote the simple closed curve $\gamma_1 + \beta_1 + \beta_2 + \gamma_2$ by J''. If any one of the conditions $c_1 = p_3$, $c_2 = p_3$, or $c_2 = p_5$ are true, the simple closed curve containing K is evident. Thus we assume that the point p_3 is an interior point of J'', and by use of Lemma 1, there are three independent arcs δ_1, δ_2, and δ_3 which join p_3 to three distinct points d_1, d_2, and d_3 respectively, of J''. We consider the possible distributions of the points d_i on the arcs $\gamma_1 = p_4c_1$, $\beta_1 = c_1p_5$, and $\gamma_2 + \beta_2 = p_4p_5$, whose sum is J''. If two of the d_i are on one of the arcs γ_1, β_1, or $\gamma_2 + \beta_2$, then the simple closed curve is again evident; e. g., when d_1 and d_2 are on γ_1 then $p_2c_3 + \beta_3 + \beta_2 + \gamma_2 + p_4d_2 + \delta_2 + \delta_1 + d_1c_1 + c_1p_1 + \alpha_2$ contains K. If no one of the arcs of J'' contains two of the d_i, then each arc contains exactly one. Assume that d_1 is on γ_1, d_2 is on β_1, and d_3 is on $\gamma_2 + \beta_2$. Then $p_2c_3 + \beta_3 + \beta_2 + \gamma_2 + p_4d_1 + \delta_1 + \delta_2 + d_2c_1 + c_1p_1 + \alpha_2$ is a simple closed curve containing K. This completes the proof for case II_b.

Case II_c. Let β_1, β_2, and β_3 be three independent arcs joining p_5 to b_1, b_2, and b_3 respectively on the boundary of D_3. Let the subarcs p_1p_3 and p_3p_2 of α_1 be α'_1 and α''_1. Similarly let p_4 divide α_2 into the subarcs α'_2 and α''_2. The simple closed curve containing K is evident if two of the points b_i fall on the same subarc of α_1 or α_2. Suppose b_1 is on α'_1, b_2 is on $\alpha''_1 - p_2$, and b_3 is on α'_2, any other combination being similar. Then $p_1p_3 + p_3b_2 + \beta_2 + \beta_3 + b_3p_4 + \alpha''_2 + \alpha_3$ is a simple closed curve containing K.

If $b_2 = p_2$, and $b_3 = p_4$, then the simple closed curve $\alpha'_2 + \beta_3 + \beta_1 + b_1p_3 + \alpha''_1 + \alpha_3$ will contain K. We now suppose that $b_2 = p_2$ and $b_3 \neq p_4$. Then p_4 is an interior point of the simple closed curve $p_1b_3 + \beta_3 + \beta_2 + \alpha_3$

---

[16] If α is an arc with end-points x and y, then by the symbol $\langle\alpha\rangle$ we shall mean $\alpha - (x + y)$.

which we denote by J'''. Then let γ_1, γ_2, and γ_3 denote the three arcs (Lemma 3) which join p_4 to distinct points c_1, c_2, and c_3, respectively, of J'''. If two of the points c_i fall on one of the arcs $p_1b_3 + \beta_3$, β_2, and α_3, their end points included, then the simple closed curve is evident. We thus assume that c_1 lies on $\langle p_1b_3 + \beta_3\rangle$, c_2 on $\langle\beta_2\rangle$, and c_3 on $\langle\alpha_3\rangle$. Then $p_1b_3 + \beta_3 + \beta_1 + b_1p_3 + \alpha''_1 + p_3c_3 + \gamma_2 + \gamma_3 + c_3p_1$ is the desired simple closed curve.

Case II_d. Here again we need discuss only the case where p_5 is in D_1. Consider the end-points b_i $(i = 1, 2, 3)$ of the independent arcs β_i from p_5 to the simple closed curve $\alpha_2 + \alpha_3$. If two of the points, say b_1 and b_2, belong to α_3, assuming the order $p_1b_1b_2p_2$ on α_3, then let $\bar{\alpha}_3$ be the arc $p_1b_1 + \beta_1 + \beta_2 + b_2p_2$. Now α_1, α_2, and $\bar{\alpha}_3$ each contain a point of K as an interior point and the conditions of Case I_b are satisfied. If two of the b_i fall on α'_2 or α''_2 of α_2, then let $\bar{\alpha}_2$ be defined similarly to $\bar{\alpha}_3$. The arc $\bar{\alpha}_2$ contains p_4 and p_5 as interior points and the simple closed curve $\bar{\alpha}_2 + \alpha_1$ contains K. If no two points b_i fall on one of the arcs α_3, α'_2, or α''_2, then each arc contains exactly one of the b_i. Suppose b_1 is on α_3, and b_2 is on α'_2, then let $J = p_1b_1 + \beta_1 + \beta_2 + b_2p_4 + \alpha''_2 + \alpha_1$. This proves the theorem for Case *II*.

Case III_a. Assume that p_4 and p_5 are in D_3, the other possibility being symmetrical. Let the three independent arcs β_i join p_4 to b_i on $\alpha_1 + \alpha_2$. If two of the b_i belong to the same subarc of α_1, then define $\bar{\alpha}_1$ as above to include p_3 and p_4. Then the situation is covered by II_b unless p_5 belongs to $\bar{\alpha}_1$. In this event let $J = \bar{\alpha}_1 + \alpha_2$. If two of the b_i belong to α_2, then define $\bar{\alpha}_2$ to include p_4 and apply II_c or II_d unless p_5 belongs to $\bar{\alpha}_2$. In this case let $J = \bar{\alpha}_2 + \alpha_1$. Finally suppose b_1 is on p_1p_3 of α_1, b_2 is on p_3p_2 of α_1, and b_3 is on α_2. If p_5 is on one of the arcs β_i, the simple closed curve which contains K is apparent.

The arcs β_i divide D_3 into three regions which we will denote by D', D'', and D'''. Since p_5 is not on the boundary of one of these regions it must be an interior point of one of them. Suppose that D' is bounded by $J' = \beta_1 + \beta_3 + b_3p_1 + p_1b_1$, D'' by $J'' = \beta_1 + \beta_2 + b_2p_3 + p_3b_1$, and D''' by $J''' = \beta_2 + \beta_3 + b_3p_2 + p_2b_2$. The discussion for p_5 in D''' is the same as that for p_5 in D'. Suppose now that p_5 is in D'. Let γ_1, γ_2, and γ_3 be three independent arcs joining p_5 to c_1, c_2, and c_3 on J'. If two of the c_i belonging to the same subarc of J', then it is evident that there is a simple closed curve containing K. In the contrary case, suppose that c_1 is on b_1p_1, c_2 is on β_1, and c_3 is on β_2. Then $p_1c_1 + \gamma_1 + \gamma_2 + c_2b_1 + b_1p_3 + p_3b_2 + \beta_2 + \beta_3 + b_3p_2 + \alpha_3$ contains the set K. For other combinations of the c_i on the four subarcs of J', the simple closed curve containing the set K can be read similarly.

For p_5 in D'', we proceed as follows. If two of the points c_i, which are again end-points of independent arcs γ_i, belong to the same subarc of J'', then J is clearly seen. If no two of the c_i belong to the same subarc of J'', then one point must be on either p_3b_1 or b_2p_3, while another one must belong to β_1 or to β_2. Suppose that c_1 is on p_3b_1, and that c_2 belongs to β_1. Then the simple closed curve $p_2b_2 + b_2p_3 + p_3c_1 + \gamma_1 + \gamma_2 + c_2p_4 + \beta_3 + b_3p_1 + \alpha_3$ contains K; and similarly for other combinations of arcs. This concludes the proof for this case.

Case III_b. For p_4 and p_5 belonging to D_1, we define $\bar{\alpha}_2$ or $\bar{\alpha}_3$ (depending on which contains at least two of the b_i) to include p_4. The case is then the same as II_c or II_d, unless $\bar{\alpha}_2$ or $\bar{\alpha}_3$ contains p_5 when the simple closed curve containing K is obvious.

Case III_c. We shall prove this case by reducing it to one of the previous cases. Let β_1, β_2, and β_3 be three independent arcs joining p_4 to $\alpha_1 + \alpha_2$, with end-points b_i. If two of the b_i belong to the same subarc p_1p_3, p_3p_2, or to α_2, then either α_1 or α_2 can be redefined to include p_4. The situation now is the same as II_b or II_d. Suppose then that p_1p_3 contains b_1, p_3p_2 contains b_2, and α_2 contains b_3. Let $\bar{\alpha}_1 = p_1b_1 + \beta_1 + \beta_2 + b_2p_2$; then α_1, α_2, and α_3 will be an independent set of arcs joining p_1 and p_2, such that p_4 belongs to $\bar{\alpha}_1$ and $p_3 + p_5$ lie together in one of the complementary domains of $\bar{\alpha}_1 + \alpha_2 + \alpha_3$. This satisfies the hypothesis of III_a. Thus case III_c is proved.

Case III_d. Let β_1, β_2, and β_3 be independent arcs joining p_5 to the boundary of $D_1 = \alpha_2 + \alpha_3$. Two of the end-points b_i of β_i lie on either α_2 or α_3. Then as above, define $\bar{\alpha}_2$ or $\bar{\alpha}_3$ to include p_5 and the case reduces to II_c or to II_d. This completes the proof for $K = 5$ points.

(NOTE. The theorem due to Ayres[17] shows that Theorem 1 is true for $K = 2$ or 3 points, and we have shown the existence of J for $K = 5$ points. For the case of $K = 4$ points we proceed exactly as in the case for 5 points but, having fewer points to consider, the analysis is much simpler.)

THEOREM 2. *If M is 3-connected, and K consists of $\leqq 5$ points, then there exists a simple closed curve J such that $M \supset J \supset K$.*

Proof. If M is 3-connected then any subset of M is 3-connected in M. Hence Theorem 2 is a direct corollary of Theorem 1.

We now show by means of several examples that the conditions imposed on Theorem 1 can not be weakened.

---

[17] *Loc. cit.*[1]

Example 1. We show here that Theorem 1 is not true for $K \geqq 6$ points. Let a_1, a_2, and a_3 be the vertices of an equilateral triangle T. Let T' be the triangle formed by joining the midpoints of the sides of T. Let x be a point exterior to T, and y be a point interior to T'. Draw independent arcs xa_i $(i = 1, 2, 3)$ and yb_i $(i = 1, 2, 3)$ where the b_i are the midpoints of the sides of T'. It is easily seen that $K = \Sigma a_i + \Sigma b_i$ is 3-connected in M and that there is no simple closed curve in M which contains the point set K. In this example the points of K are all of order 3, while the points of M are all of order $\leqq 4$.

Example 2. In this example we wish to show that Theorem 2 cannot be extended to include more than 5 points in K. Let a circle C be inscribed in an equilateral triangle T with vertices x_1, x_2, and x_3. Let the points of tangency be p_1, p_2, and p_3. Denote the triangular domains of $T - C$ by C_1, C_2, and C_3. In the circle C construct three independent arcs α_1, α_2, and α_3 with end-points p_1 and p_2, p_2 and p_3, p_3 and p_1 respectively. In the domain D whose boundary is $\Sigma\alpha_i$, construct three arcs β_1, β_2, and β_3 which are distinct except for their end-points p_1 and p_2, p_2 and p_3, p_3 and p_1. Denote the region between α_i and β_i by D_i $(i = 1, 2, 3)$.

Let b_i denote the midpoints of the arcs β_i, and construct distinct arcs b_ib_{i+1} (mod 3) in the interior of the domain bounded by $\Sigma\beta_i$. Let B denote the domain whose boundary is Σb_ib_{i+1}. Then we let M denote the point set $\Sigma\bar{C}_i + \Sigma\bar{D}_i + \bar{B}$ where the vertices x_1, x_2, and x_3 are now identified. If c_i is a point of C_i and d_i is a point of D_i, then let K be the six points $\Sigma c_i + \Sigma d_i$. It is evident that M is 3-connected and that every point of M is of the order of the continuum, but it is not possible for K to be included in a simple closed curve in the set M. This example also shows that a previously announced theorem[18] by Ayres is not true for $K > 5$ points.

Example 3. The following example[19] shows that if the condition that M be planar is removed, then Theorem 1 is not true even for K consisting of but four points. In the Euclidean 3-dimensional space, let p_i $(i = 1, \cdots, 4)$ be the points $(1, 0, 0)$, $(0, 0, 1)$, $(0, 0, -1)$, and $(-1, 0, 0)$, and let q_j $(j = 1, 2, 3)$ be the points $(0, -1, 0)$, $(0, 0, 0)$, and $(0, 1, 0)$. Let $M = \Sigma p_iq_j$, where p_iq_j denotes the straight line interval from p_i to q_j, and let $K = \Sigma p_i$. It is easy to see that K is 3-connected in M, and cannot be joined by a simple closed curve in M.

---

[18] The statement of the theorem is the same as that of Theorem 2 except that K is *any* finite set. This result was announced in an abstract No. 51, p. 772 *Bulletin of the American Mathematical Society*, vol. 35 (1929), but never published as an error was found in the proof.

[19] Ayres, *loc. cit.*[1]

Example 4. Suppose we allow K to be only 2-connected in M, by taking M to be the θ-curve consisting of the three independent arcs xa_iy and let $K = \Sigma a_i$. Then K cannot be joined by a simple closed curve in M.

Example 5. We show here that if M is not a Peano space, even though it be arc-wise[20] connected, and K is 3-connected in M and consists of only 2 points, then K may not be included in a simple closed curve in M. In the Euclidean plane, let S denote the curve $y = \sin \pi/x$ $(0 < x \leqq 1)$. We add the limit points and an arc α which joins the points $(1, 0)$ and $(0, 1)$ but otherwise is distinct from S. Let $M = \bar{S} + \alpha$, and let K be the two points $(0, 1)$ and $(0, -1)$. We observe that K is 3-connected in M, and that M is arc-wise connected, but contains no simple closed curve whatever. Therefore K is not contained in a simple closed curve. This example shows further that Rutt's theorem[21] is not true if M is not a Peano space, even though it be arc-wise connected.

THEOREM 3. *For any given positive number* $m \geqq 3$, *there exists a planar Peano space* M *containing a set* $K = 2m$ *points which is 3-connected in* M *and cannot be included in* $m - 2$ *mutually exclusive simple closed curves in* M.

The set M which we shall construct is a generalization of Example 1, in which case $m = 3$. Let N be a regular m-gon with vertices $a_1, a_2, \cdots, a_m$. Let N' be the regular m-gon formed from N by joining the midpoints of adjacent sides by lines. Denote the midpoints of the sides of N' by $b_1, b_2, \cdots, b_m$ and the vertices of N' by $z_1, z_2, \cdots, z_m$. Let y be an interior point of N' and x be a point exterior to N. Draw m independent arcs xa_i and m independent arcs yb_i. Let $M = \Sigma xa_i + \Sigma yb_i + N + N'$ and let $K = \Sigma a_i + \Sigma b_i$.

We shall now show that K can not be included in $m - 2$ mutually exclusive simple closed curves. From the construction it is evident that every simple closed curve in M which contains s points of K will also contain s points of the set $Z = x + y + \Sigma z_i$. Since there are only $m + 2$ points in Z, then, irrespective of the manner in which the mutually exclusive simple closed curves are constructed, they can contain at most $m + 2$ points of K. There remain then at least $m - 2$ isolated points of K which can be included only in degenerate simple closed curves. Assuming the most favorable case, namely that $m + 2$ points are contained in a simple closed curve (this can be achieved), then $m - 1$ mutually exclusive simple closed curves are necessary to contain K.

THE COLLEGE OF WOOSTER,
WOOSTER, OHIO.

---

[20] A point set M is *arc-wise* connected when every two points of M are the extremities of a simple continuous arc lying wholly in M.

[21] *Loc. cit.*[9]

THE EXPANSION OF AN ARBITRARY FUNCTION IN A SERIES OF TOROIDAL FUNCTIONS OF THE SECOND KIND.*

By D. P. BANERJEE.

Recently[1] I have given the theorem of expansion of an arbitrary function in a series of toroidal functions of the first kind. These functions are required in electrostatical and hydrodynamical problems connected with an anchor-ring. I have also considered the expansion in a series of associated toroidal functions of the second kind (also a new and interesting integral property of toroidal functions). Prasad[2] and N. G. Shabde[3] have summed certain similar series without actually giving the process of summation of an arbitrary function.

We know[4] that

$$Q_n(z) = \int_0^{z-(z^2-1)^{\frac{1}{2}}} \frac{h^n\,dh}{(1-2zh+h^2)^{\frac{1}{2}}}$$

if $R(n+1) > 0$ and $|z| > 1$.

Then

$$Q_{n+\frac{1}{2}}(z) = \int_0^{z-(z^2-1)^{\frac{1}{2}}} \frac{h^{n+\frac{1}{2}}\,dh}{(1-2zh+h^2)^{\frac{1}{2}}}$$

$Q_{n+\frac{1}{2}}(z)$ being the toroidal function of the second kind. Let

$$f(z) = \sum_0^{\infty} a_n\, Q_{n+\frac{1}{2}}(z).$$

In order to find a_n let

$$f(z) = \int_0^{z-(z^2-1)^{\frac{1}{2}}} \frac{\phi(h^{\frac{1}{2}})\,dh}{(1-2zh+h^2)^{\frac{1}{2}}}$$

where $\phi(h)$ is an odd analytic function of h. Then

$$f(z) = \int_0^{z-(z^2-1)^{\frac{1}{2}}} \frac{\phi(h^{\frac{1}{2}})\,dh}{(2h)^{\frac{1}{2}}\left(\frac{1+h^2}{2h} - z\right)^{\frac{1}{2}}}$$

---

* Received June 5, 1939; revised February 27, 1940.

[1] D. P. Banerjee, "The Expansion of an arbitrary function in a series of conal or toroidal functions," *Proceedings of the Cambridge Philosophical Society*, vol. 34, pt. 1, p. 30.

[2] G. Prasad, *Bulletin Calcutta Mathematical Society*, vol. 23 (1931), pp. 155-182.

[3] N. G. Shabde, *Bulletin Calcutta Mathematical Society*, vol. 23 (1931), pp. 23-44.

[4] *Modern Analysis*, 3rd Ed. (Cambridge), by Whittaker and Watson, p. 334.

Let $h_1 = \frac{2h}{1+h^2}$ so that $dh_1 = \frac{2(1-h^2)}{(1+h^2)^2}\,dh$ and let

$$z = \frac{1}{z_1};\ \phi_1'(h_1^{\frac{1}{2}}) = \frac{\phi(h^{\frac{1}{2}})\ (1+h^2)^2}{2(1-h)^2}\left(\frac{h_1}{2h}\right)^{\frac{1}{2}}$$

Then

$$\frac{f\left(\frac{1}{z}\right)}{\sqrt{z_1}} = \int_0^{z_1} \frac{\phi'_1((h_1)^{\frac{1}{2}})\ dh_1}{(z_1-h_1)^{\frac{1}{2}}}\ ;\ |\,z_1\,| < 1$$

Let $h_1 = z_1 \sin^2\theta$; and let $F(z_1) = \frac{f\left(\frac{1}{z_1}\right)}{\sqrt{z_1}}$; then

$$F(z_1) = \int_0^{\pi/2} \phi'_1(z_1^{\frac{1}{2}}\sin\theta)\,2z_1^{\frac{1}{2}}\sin\theta d\theta;$$

$$z_1\int_0^{\pi/2}\frac{F(z_1^2\sin^2\psi)\,d\psi}{2z_1\sin\psi} = z_1\int_0^{\pi/2}\left\{\int_0^{\pi/2}\sin\theta\phi'_1(z_1\sin\theta\sin\psi)\,d\theta\right\}d\psi.$$

On putting $\sin\chi = \sin\psi\sin\theta$ and changing the order of integration we have

$$z_1\int_0^{\pi/2}\frac{F(z_1^2\sin^2\psi)\,d\psi}{2z_1\sin\psi} = z_1\int_0^{\pi/2}\int_\chi^{\pi/2}\frac{\phi'_1(z_1\sin\psi)\ \cos\psi\sin\theta\ d\theta d\psi}{(\sin^2\theta-\sin^2\psi)^{\frac{1}{2}}}$$
$$= \frac{\pi}{2}\,[\phi_1(z_1)-\phi_1(0)].$$

Hence

$$(1)\quad \phi'_1(z_1) = \frac{2}{\pi}\left[\int_0^{\pi/2}\frac{F(z_1^2\sin^2\psi)\,d\psi}{2z_1\sin\psi} + z_1\frac{d}{dz_1}\int_0^{\pi/2}\frac{F(z_1^2\sin^2\psi)\,d\psi}{2z_1\sin\psi}\right]$$

and

$$\phi(h^{\frac{1}{2}}) = \frac{2(1-h^2)}{(1+h^2)^{3/2}}\ \phi'_1(h_1^{\frac{1}{2}})$$

Example. Let

$$f(z) = \frac{1}{z^{m+\frac{1}{2}}};\ \text{then } F(z_1) = z_1^{\,m}$$

$$\phi'_1(z_1) = \frac{1}{\pi}\left[z_1^{\,2m-1}\int_0^{\pi/2}\sin^{2m-1}\psi d\psi + (2m-1)z_1^{\,2m-1}\int_0^{\pi/2}\sin^{2m-1}\psi d\psi\right]$$
$$= \frac{1}{\pi}\cdot\ \frac{\Gamma(m+1)\Gamma(\frac{1}{2})}{\Gamma(m+\frac{1}{2})}\,z_1^{\,2m-1}$$

$$\phi(h^{\frac{1}{2}}) = \frac{2^{m+\frac{1}{2}}}{\pi}\,\frac{\Gamma(m+1)\Gamma(\frac{1}{2})}{\Gamma(m+\frac{1}{2})}\,h^{m-\frac{1}{2}}\left[1-(m+2)h^2+\frac{(m+1)(m+4)}{2}h^4-\cdots\right]$$

Hence

$$\frac{1}{z^{m+\frac{1}{2}}} = \frac{2^{m+\frac{1}{2}}}{\pi}\,\frac{\Gamma(m+1)\Gamma(\frac{1}{2})}{\Gamma(m+\frac{1}{2})}\left[Q_{m-\frac{1}{2}}(z)-(m+2)\,Q_{m+3/2}(z)\right.$$
$$\left.+\frac{(m+1)(m+4)}{2}\,Q_{m+7/2}(z)-\cdots\right]$$

If $m = 1$

$$\frac{1}{z^{3/2}} = \frac{2^{5/2}}{\pi}\left[Q_{1/2}(z) - 3Q_{5/2}(z) + 5Q_{9/7}(z) - \cdots\right].$$

The Test of Convergency:—

since

$$Q_{n+\frac{1}{2}}(z) = 0 \text{ at } z = \infty$$

$f(z)$ must be zero at $z = \infty$ and must be analytic when $|z| > 1$. Since $|z| > 1$, $|z_1| < 1 < \pi$ and consequently the solution of the integral equation (1) is valid.[5] Let $F(z_1)$ be an analytic function of z_1 which can be expanded in Maclaurin's series. Now the series

$$F(z_1) = \sum_0^\infty \frac{F^n(0)}{n!} z_1^n$$

will certainly be convergent if

$$\underset{n\to\infty}{Lt} \frac{(n+1)F^n(0)}{F^{n+1}(0)z_1} = 0$$

But

$$\frac{\pi}{2}\phi'_1(z_1) = \sum_1^\infty \frac{F^n(0)}{(n-1)!} z_1^{2n-1} \int_0^{\pi/2} \sin^{2n-1}\psi \, d\psi = \sum_1^\infty \frac{\Gamma(\frac{1}{2})}{\Gamma(n+\frac{1}{2})} F^n(0) z_1^{2n-1}$$

so that $\phi'_1(z_1)$ will be convergent if

$$\underset{n\to\infty}{Lt} \frac{F^{n+1}(0)}{F^n(0)} \frac{z_1^2}{n+\frac{1}{2}} = 0,$$

or, if

$$\underset{n\to\infty}{Lt} \frac{F^{n+1}(0)z_1^2}{(n+1)F^n(0)\left[1 - \dfrac{1}{2(n+1)}\right]} = 0$$

when $F(z_1^2)$ or $F(z_1)$ are analytic. Since $F(z_1)$ is an analytic function which can be expanded in a convergent Maclaurin's series we know that

$$\phi(z_1) = \frac{2(1-z_1^2)}{(1+z_1^2)^{3/2}} \phi_1(z_1)$$

can be expanded in a convergent power series. Hence the series in toroidal functions of the second kind will certainly be convergent when

(i) $|z| > 1$

(ii) $F(z_1)$ is an analytic function of z_1 which can be expanded in a convergent power series.

---

[5] Whittaker and Watson, *loc. cit.*, p. 229.

(iii) $F(0) = 0$

(iv) $\underset{z\to\infty}{Lt}\, f(z) = 0$ and $f(z)$ is a convergent analytic function when $|z| > 1$

Singularities of $f(z)$.

(a) Let $f(z) = z^{-\frac{1}{2}}$, then $F(z_1) = 1$ and consequently the previous method fails. Now from (2) we have

$$\frac{1}{z^{3/2}} = \frac{2^{5/2}}{\pi}\left[Q_{1/2}(z) - 3Q_{5/2}(z) + 5Q_{9/2}(z) - \cdots\right].$$

Hence

$$z^{-1/2} = \frac{2^{5/2}}{\pi} z\left[Q_{1/2}(z) - 3Q_{5/2}(z) + 5Q_{9/2}(z) - \cdots\right].$$

Now applying the formulae

$$(n + 1/2)\, Q_{n+1/2}(z) + (n - 1/2)\, Q_{n-3/2}(z) = 2nz\, Q_{n-1/2}(z).$$

We have

$$z^{-1/2} = \frac{2^{3/2}}{\pi}\left[\tfrac{1}{2}Q_{-1/2}(z) - Q_{3/2}(z) + Q_{7/2}(z) - \cdots\right]$$

(b) Let $f(z) = 1/z^p$ where $0 < p < \frac{1}{2}$. Then $F(z_1) = z_1^{p-1/2}$ and $F(0) = \infty$ and consequently the previous method cannot be applied. Let $f_1(z) = 1/z^{p+1}$. Then $F_1(z_1) = z_1^{p+1/2}$ and $F_1(0) = 0$ and we can apply the previous method. Let

$$f_1(z) = \sum_0^\infty a_n Q_{n+1/2}(z)$$

where the a_n's are found by the previous method. Then

$$f(z) = z\sum_0^\infty a_n Q_{n+1/2}(z) = \sum_0^\infty a_n' Q_{n-1/2}(z)$$

where we get the a'_n's from the a_n's by applying the recurrence formulae.

II.

The associated toroidal function of the second kind is defined by

$$Q^m{}_{n+1/2}(z) = (z^2 - 1)^{m/2}\frac{d^m}{dz^m} Q_{n+1/2}(z),$$

Let

$$F(z) = \sum_0^\infty a_n Q^m{}_{n+1/2}(z) = (z^2 - 1)^{m/2}\frac{d^m}{dz^m} f(z) = (z^2 - 1)^{m/2}\frac{d^m}{dz^m}\sum_0^\infty a_n Q_{n+1/2}(z).$$

Then

$$f(z) = \int_0^z dz \int_0^z dz \cdots \int_0^z \frac{F(z)\,dz}{(z^2-1)^{m/2}}.$$

Also

$$f(z) = \sum_0^\infty a_n Q_{n+1/2}(z)$$

so that the a_n's may be determined by the previous method.

Example.

$$F(z) = (-1)^m \cdot \frac{3 \cdot 5 \cdots (2m+1)}{2^m} \frac{(z^2-1)^{m/2}}{z^{m+3/2}}.$$

Then

$$f(z) = \frac{1}{z^{3/2}} = \frac{2^{5/2}}{\pi} [Q_{1/2}(z) - 3Q_{5/2}(z) + 5Q_{9/2}(z) - \cdots].$$

Hence

$$(-1)^m \frac{3 \cdot 5 \cdots (2m+1)}{2^m} \cdot \frac{(z^2-1)^{m/2}}{z^{m+3/2}}$$
$$= \frac{2^{5/2}}{\pi} [Q^m{}_{1/2}(z) - 3Q^m{}_{5/2}(z) + 5Q^m{}_{9/2}(z) - \cdots].$$

The test of convergency is as before.

III.

We assume

$$0 \leq z < 1. \quad \text{Let } z = \cos\theta,\ 0 < \theta \leq \frac{\pi}{2}.$$

We know[6] that

$$Q_{n+\frac{1}{2}}(-\cos\theta) = \frac{\pi}{2} \ (-1)^{n+1} P_{n+\frac{1}{2}}(\cos\theta).$$

$$P_{n+\frac{1}{2}}(-\cos\theta) = \frac{2}{\pi} \ (-1)^{n+1} P_{n+\frac{1}{2}}(\cos\theta).$$

Let

$$f(\cos\theta) = \sum_0^\infty a_n Q_{n-1/2}(\cos\theta)$$

Then

$$f(-\cos\theta) = \sum_0^\infty a_n Q_{n-\frac{1}{2}}(-\cos\theta) = \frac{\pi}{2} \sum_0^\infty (-1)^{n+1} a_n P_{n+\frac{1}{2}}(\cos\theta)$$

Let

$$F(\cos\theta) = \frac{2}{\pi} f(-\cos\theta) = \sum_0^\infty b_n P_{n-\frac{1}{2}}(\cos\theta).$$

Then

$$a_n = (-1)^{n+1} 6n.$$

---

[6] G. Prasad, "A Treatise on Spherical Harmonics and the functions of Bessel and Lamé," Part II, p. 61.

Now[7]

$$P_{n-\frac{1}{2}}(\cos\theta) = \frac{2}{\pi}\int_0^\theta \frac{\cos n\phi\, d\phi}{[2(\cos\phi - \cos\theta)]^{\frac{1}{2}}}$$

Let

$$F(\cos\theta) = \frac{2}{\pi}\int_0^\theta \frac{u(\cos\phi)\sin\phi d\phi}{[2(\cos\phi - \cos\theta)]^{\frac{1}{2}}}$$

so that

$$F(z) = \frac{2}{\pi}\int_z^1 \frac{u(\xi)\,d\xi}{[2(\xi - z)]^{\frac{1}{2}}}$$

Therefore by the theory of Abel's integral equation[8] we have

$$u(z) = \frac{1}{\pi}\,\frac{d}{dz}\int_z^1 \frac{F(x)\,dx}{(x-z)^{\frac{1}{2}}}\cdot\frac{\pi}{2^{\frac{1}{2}}} = \frac{1}{2^{\frac{1}{2}}}\,\frac{d}{dz}\int_z^1 \frac{F(x)\,dx}{(x-z)^{\frac{1}{2}}}$$

We note that $F(1) = 0$ and assume $F'(x)$ to be continuous over

$$0 \leq \phi \leq \theta \leq \frac{\pi}{2}.$$

Let

$$[u(\cos\phi) - u(-\cos\phi)]\sin\phi = 2\sum_0^\infty b_{2n}\cos 2n\phi.$$

and

$$[u(\cos\phi) + u(-\cos\phi)\sin\phi = 2\sum_0^\infty b_{2n+1}\cos(2n+1)\,\phi.$$

Then

$$u(\cos\phi)\sin\phi = \sum_0^\infty b_n\cos n\phi.$$

Now

$$\int_0^{\pi/2}\cos^2 m\phi d\phi = \frac{\pi}{4}$$

$$b_{2n+1} = \frac{2}{\pi}\int_0^{\pi/2}[u(\cos\phi) + u(-\cos\phi)]\sin\phi\cos(2n+1)\phi\, d\phi$$

$$b_{2n} = \frac{2}{\pi}\int_0^{\pi/2}[u(\cos\phi) - u(-\cos\phi)]\sin\phi\cos 2n\phi d\phi.$$

$$b_0 = \frac{1}{\pi}\int_0^{\pi/2}[u(\cos\phi) - u(-\cos\phi)]\sin\phi d\phi$$

Hence

$$a_0 = \frac{1}{\pi}\int_0^{\pi/2}[u(\cos\phi) - u(-\cos\phi)]\sin\phi d\phi.$$

$$a_{2n} = \frac{2}{\pi}\int_0^{\pi/2}[u(\cos\phi) - u(-\cos\phi)]\sin\phi\cos 2n\phi d\phi$$

$$a_{2n+1} = -\frac{2}{\pi}\int_0^{\pi/2}[u(\cos\phi) + u(-\cos\phi)]\sin\phi\cos(2n+1)\phi d\phi.$$

---

[7] Hobson, *The Theory of spherical and ellipsoidal Harmonics* (Cambridge, 1931), p. 267, equation (130).

[8] Whittaker and Watson, *loc. cit.*, p. 229.

The test of convergency depends upon

(1) The convergency of the corresponding Fourier's series.

(2) The existence of the solution of the Abel integral equation.

Example. Let

$$f(\cos\theta) = \frac{2}{\pi}\int_0^{\pi/2} \frac{\phi \sin\phi d\phi}{[2(\cos\phi + \cos\theta)]^{\frac{1}{2}}}$$

Then

$$f(z) = \frac{2}{\pi}\int_{-z}^{1} \frac{\cos^{-1}\xi d\xi}{[2(\xi - z)]^{\frac{1}{2}}}$$

$$f(-z) = \frac{2}{\pi}\int_{z}^{1} \frac{\cos^{-1}\xi d\xi}{[2(\xi - z)]^{\frac{1}{2}}}$$

$$f(-\cos\theta) = \frac{2}{\pi}\int_0^{\theta} \frac{\phi \sin\phi d\phi}{[2(\cos\phi + \cos\theta)]^{\frac{1}{2}}}$$

$$u(\cos\phi) = \phi;\ 0 \leq \phi < \frac{\pi}{2}.$$

$$\phi\sin\phi = \sum_0^\infty b_{2n+1}\cos(2n+1)\phi + \sum_0^\infty b_{2n}\cos 2n\phi$$

$$b_{2n+1} = \frac{4}{\pi}\int_0^{\pi/2}\phi\sin\phi\cos(2n+1)\phi d\phi = \frac{(-1)^n(2n+1)}{2n(n+1)};$$

$$b_{2n} = 0$$

$$f(-\cos\theta) = \sum_0^\infty \frac{(-1)^n(2n+1)}{2n(n+1)} P_{2n+\frac{1}{2}}(\cos\theta)$$

$$f(\cos\theta) = -\frac{1}{\pi}\sum_0^\infty \frac{(2n+1)}{n(n+1)} Q_{2n+\frac{1}{2}}(\cos\theta).$$

Therefore

$$\frac{2}{\pi}\int_0^{\pi-\theta} \frac{\phi\sin\phi d\phi}{[2(\cos\phi + \cos\theta]} = -\frac{1}{\pi}\sum_0^\infty \frac{2n+1}{n(n+1)} Q_{2n+\frac{1}{2}}(\cos\theta)$$

for $0 < \theta \leq \frac{\pi}{2}$.

IV.

If $-1 < z < 1$ and $n - m \neq 0$; $n + m + 1 \neq 0$

$$(m-n)(m+n+1)\int_{-1}^{1} P_n(z) Q_m(z) dz$$

$$= -1 + \cos(m-n)\pi - \frac{2}{\pi}\sin\pi n\cos m\pi\,[\psi(m) - \psi(n)];$$

m, n being unrestricted.

From the differential equation and recurrence formulae satisfied by $P_n(z)$, $Q_m(z)$ and using the formulae[9]

---

[9] G. Prasad, *loc. cit.*, pp. 61, 72.

$$P_n(-\cos\theta) = P_n(\cos\theta)\cos n\pi - \frac{2}{\pi}\sin n\pi\, Q_n(\cos\theta)$$

$$Q_n(-\cos\theta) = -Q_n(\cos\theta)\cos n\pi - \frac{\pi}{2}\sin n\pi\, P_n(\cos\theta)$$

$$\begin{aligned} Q_n(z) = {} & \frac{1}{2}P_n(z)\log\frac{1+z}{1-z} + \left\{\pi'(0) - \frac{\pi'(n)}{\pi(n)}\right\}P_n(z) \\ & - \sum_1^\infty \frac{\pi(-n+r-1)\pi(n+r)}{\pi(n)\pi(-n-1)\{\pi(r)\}^2} A_r \left(\frac{1-z}{2}\right)^r \end{aligned}$$

where

$$A_r = \sum_1^r \frac{1}{k}\;.$$

We have

$$\begin{aligned} & (m-n)(m+n+1)\int_{-1}^{1} P_n(z)Q_m(z)\,dz \\ & = {}_{-1}^{\;1}\big[z(m-n)P_n(z)Q_m(z) - mP_n(z)Q_{m-1}(z) + nQ_m(z)P_{n-1}(z)\big] \\ & = m[Q_m(z) - Q_{m-1}(z)]_{z=1} + mP_m(-z)[Q_m(-z) + Q_{m-1}(-z)]_{z=1} \\ & \quad - nQ_m(-z)[P_n(-z) + P_{n-1}(-z)]_{z=1}\;. \end{aligned}$$

Now

$$Q_m(z) - Q_{m-1}(z) = \psi(m-1) - \psi(m) = -\frac{1}{m} \text{ when } z = 1$$

$$\begin{aligned} Q_m(-z) + Q_{m-1}(-z) & = -\cos m\pi[Q_m(z) - Q_{m-1}(z)] \\ & = \frac{\cos m\pi}{m}\;. \end{aligned}$$

When $z = 1$

$$\begin{aligned} P_n(-z) + P_{n-1}(-z) & = -\frac{2}{\pi}\sin n\pi[Q_n(z) - Q_{n-1}(z)]_{z=1} \\ & = \frac{2}{\pi}\frac{\sin n\pi}{n}\;. \end{aligned}$$

Hence

$$(m-n)(m+n+1)\int_{-1}^{1} P_n(z)Q_m(z)\,dz$$

$$\begin{aligned} & = \left[-1 + \cos m\pi\, P_n(-z) - \frac{2}{\pi}\sin n\pi\, Q_m(-z)\right]_{z=1} \\ & = \left[-1 + \cos m\pi\left\{\cos n\pi - \frac{2}{\pi}\sin n\pi\, Q_n(z)\right\} - \frac{2}{\pi}\sin n\pi\, Q_m(-z)\right]_{z=1} \\ & = \left[-1 + \cos m\pi\cos n\pi - \frac{2}{\pi}\sin n\pi\,\{Q_n(z)\cos m\pi + Q_m(-z)\}\right]_{z=1} \\ & = \left[-1 + \cos m\pi\cos n\pi - \frac{2}{\pi}\sin n\pi\,\left\{Q_n(z)\cos m\pi - Q_m(z)\cos m\pi - \frac{\pi}{2}\sin m\pi\right\}\right]_{z=1} \\ & = -1 + \cos(m-n)\pi - \frac{2}{\pi}\sin n\pi\cos m\pi\,[\psi(m) - \psi(n)] \end{aligned}$$

where

$$\psi(m) = \frac{\pi'(m)}{\pi(m)}\;.$$

Hence

$$\int_{-1}^{1} P_n(z)\, Q_m(z)\, dz = 0$$

if $m - n =$ an even integer and $m = r - \frac{1}{2}$, $n = p + \frac{1}{2}$ where r and p are positive integers and $p \neq r$.

If $R(n+1) > 0$ we know[10] that

$$(2n-1) \int_{0}^{1} P_{n-1}(x)\, Q_{n-1}(x)\, dx = (2n+1) \int_{0}^{1} P_n(x)\, Q_n(x)\, dx - \frac{\cos \pi n}{n}.$$

Hence

$$(n+1) \int_{0}^{1} P_{n+\frac{1}{2}}(x)\; Q_{n+\frac{1}{2}}(x)\, dx - n \int_{0}^{1} P_{n-\frac{1}{2}}(x)\; Q_{n-\frac{1}{2}}(x)\, dx = E$$

where

$$E = \int_{0}^{1} P_{\frac{1}{2}}(x)\; Q_{\frac{1}{2}}(x)\, dx.$$

A. M. College, Mymensingh,
Bengal, India.

---

[10] E. W. Barnes, *Quarterly Journal of Mathematics*, vol. 39 (1908), p. 204.

CHARACTERS OF ABELIAN GROUPS.*[1]

By PAUL E. LEWIS.

Introduction. The classical definition[2] of a character of a finite abelian group A is a multiplicative mapping of the group A into the multiplicative group of all the roots of unity. If the usual definition of the product of two characters is given, the set of all characters forms an abelian group $C(A)$. One of the main results in this connection is the fact that A is isomorphic to its character group. In notation: $C(A) \sim A$.

A more comprehensive theory of the characters of abelian groups which applies to every group [finite or infinite] has been constructed by Pontrjagin.[3] In order to effect this generalization, however, Pontrjagin was forced to change the system of values of the characters from the group of roots of unity to the additive group of real numbers modulo 1.

Thus there arises quite generally the question, "Which groups and which value groups will give rise to a satisfactory theory?" In order to state the problem more precisely, let A and V be any two [additively written] abelian groups. As a character one may consider any additive function[4] of the elements in A with values in V. If addition of characters is defined in the obvious fashion, then the set of all characters of A in V forms an abelian group $C_V(A)$. Our problem now is to find necessary and sufficient conditions on the pair of groups A and V such that one of the following relations holds true:

a). $C_V(A) \sim A$,
b). $C_V(S) \sim S$ for every subgroup S of A,
c). $C_V(A/S) \sim A/S$ for every subgroup S of A.

The principal results concerning this problem are special cases of a more general theory and are obtained in Section 8. Considering that the study of

---

* Received May 16, 1940; revised May 19, 1941.

[1] Presented to the Society, April 13, 1940.

[2] See, e. g., E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*, Leipzig (1923).

[3] Cf. L. Pontrjagin, "The Theory of Topological Commutative Groups," *Annals of Mathematics*, vol. 35 (1934), pp. 361-368; E. R. van Kampen, "Locally Bicompact Abelian Groups and their Character Groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.

[4] Cf., e. g., P. Alexandroff and H. Hopf, *Topologie*, Berlin (1935), p. 586; H. Whitney, "Tensor Products of Abelian Groups," *Duke Mathematical Journal*, vol. 4 (1938), pp. 495-528.

linear forms over a projective geometry with coördinates from a field (not necessarily commutative) shows all the essential features of such a theory of characters, it has seemed advisable to build up the whole theory immediately for abelian groups admitting a (not necessarily commutative) ring of operators. Among the restrictions imposed on the ring R one may mention that R must contain an identity element, and every left ideal must be a principal left ideal. This last restriction is necessary to insure that subgroups of cyclic groups will also be cyclic.

Before the theory of characters under this more general set-up is considered, it is necessary to state generalizations of many of the theorems concerning the structure of abelian groups admitting a ring R of operators. The first three sections are devoted entirely to such a development which is of considerable interest on its own account. The theory of characters is considered in the remainder of the paper, the main results being found in the last section.

1. Properties of the ring of operators.[5]

(A). *Let R be a (non-commutative) principal ideal domain, i. e., R is a domain of integrity in which every right ideal is a principal right ideal and every left ideal is a principal left ideal.*

(B). *Let every ideal in R be a two-sided ideal.*

As examples of such rings R one might mention quasi-fields and commutative principal ideal rings. Another example may be constructed as follows:[6] Let F be a field [whether commutative or not], and let f be an isomorphism of F mapping F upon a subfield F^f of F where F^f need not be equal to F. Denote by R the system of all the sequences $a = (a_0, a_1, \cdots)$ with coördinates a_i in F. If b and c are two elements in R, then $b = c$ if, and only if, $b_i = c_i$ for $i = 0, 1, 2, \cdots$; $b \pm c = (b_0 \pm c_0, b_1 \pm c_1, \cdots, b_i \pm c_i, \cdots)$; $bc = (b_0c_0, b_1c_0^f + b_0c_1, \cdots, \sum_{j+k=i} b_jc_k^{f^j}, \cdots)$. With these definitions one verifies readily that R is a ring whose null-element is $0 = (0, \cdots, 0, \cdots)$ and whose identity-element is $1 = (1, 0, 0, \cdots, 0, \cdots)$. This ring does not contain divisors of zero. Let $e(i)$ be the element in R all of whose coördinates excepting the i-th are zero and whose i-th coördinate is 1. Then one may prove that for given b the equation $xb = e(i)$ has a solution x in R if, and only if, not every $b_j = 0$ for $j \leq i$. Denoting by P_i the set of all those elements b in R satisfying $b_j = 0$

---

[5] These hypotheses for the ring R conform to those used by Asano, Nakayama, and others. See, e. g., K. Asano, "Über verallgemeinerte Abelsche Gruppen mit hyperkomplexem Operatorenring und ihre Andwendungen," *Japanese Journal of Mathematics*, vol. 15 (1939), pp. 231-253.

[6] R. Baer suggested this example.

for $j < i$, it follows from the above remark that the only left-ideals in R are these P_i, that every P_i is a two-sided ideal and that $P_i = P_1{}^i$ for every i. This implies, in particular, that the P_i are principal left ideals which may be generated by the elements $e(i)$. It can also be proved that R is commutative if, and only if, $f = 1$ and F is commutative; also every right ideal is two-sided if, and only if, $F = F^f$.

Under the above hypotheses (A) and (B) the following statements are easily derived or are already in the literature:[7]

1.1. *The multiplication of left [or right] ideals is commutative.*

1.2. *If* (s) *and* (t) *are principal left ideals, then* $(st) = (s)(t)$ *for every* s, t *in* R.

1.3. R *is a field if, and only if,* 0 *and* R *are its only ideals.*

A *prime* left ideal is defined to be a greatest left ideal in R that is different from both 0 and R. Since in R divisors are factors, this is equivalent to the more common definition that a prime left ideal is one that contains a product if, and only if, it contains one of the factors.

1.4. *If* (p) *is a prime left ideal in* R, *then* $R/(p)$ *is a field.*

1.5. *Every left ideal is a product of a finite number of prime left ideals.*

1.6. *Two products of prime left ideals are equal if, and only if, they are products of the same prime left ideals apart from order of factors.*

2. Groups admitting R as operator-system.[8] Consider an abelian group A (additively written) admitting the ring R as a system of left-operators. The set of elements $B = Rx$ formed by the totality of elements rx for r in R forms a subgroup of A. B is called the cyclic subgroup of A generated by the element x: $B = Rx = \{x\}$. The order of an element x in A is defined to be the set $O(x)$ of all elements w in R such that $wr = 0$. It is easy to show that $O(x)$ is a left ideal in R. Moreover, the order of every element t in $B = \{x\}$ divides the order of x: $O(t) \geq O(x)$. From this statement it follows that all the generating elements of a cyclic group have the same order and this order ideal is defined to be the order of the cyclic group. Condition (A) implies that every subgroup of a cyclic group is also cyclic.

---

[7] T. Nakayama, "A note on elementary divisor theory in a non-commutative domain," *Bulletin of the American Mathematical Society*, vol. 44 (1938), pp. 719-723.

[8] In this connection one should mention G. Kothe, "Verallgemeinerte Abelsche Gruppen mit hyperkomplexem Operatorenring," *Math. Zeitschrift*, vol. 39 (1935), pp. 31-44. See also footnote 4.

Two groups A and B are said to be isomorphic [in notation: $A \sim B$] if there exists a $(1-1)$ mapping α of the elements of A on the elements of B such that if x, y are in A, then $(x+y)^\alpha = x^\alpha + y^\alpha$; $(rx)^\alpha = r \cdot x^\alpha$ for r in R.

2.1. *Let Z and Y be cyclic groups. Then $O(Z) = O(Y)$ if, and only if, $Z \sim Y$.*

Proof. Assume $Z \sim Y$, that is $Z^\beta = Y$ for some isomorphism β. Let $Z = \{z\}$ and define $y = z^\beta$ so that $Y = Z^\beta = \{z\}^\beta = \{y\}$. Then $rz = 0$ implies $0 = 0^\beta = (rz)^\beta = rz^\beta = ry$. Hence $O(z) \leq O(y)$. Similarly $O(y) \leq O(z)$, and this proves $O(Z) = O(Y)$. Conversely suppose $O(Z) = O(Y)$ where $Z = \{z\}$, and $Y = \{y\}$. Now $rz = sz$ if, and only if, $ry = sy$. For $rz = sz$ implies $(r-s)z = 0$, so that $r-s$ is in $O(z)$. But $O(Z) = O(Y)$, and hence $(r-s)y = 0$ or $ry = sy$. The converse of this statement follows in the same way. Now define $(rz)^\beta = ry$. From the above statement it follows that β is a single-valued function. Furthermore

$$\begin{aligned}(sz+tz)^\beta = ([s+t]z)^\beta = (s+t)z^\beta &= (s+t)y \\ &= sy+ty = s(z)^\beta + t(z)^\beta = (sz)^\beta + (tz)^\beta;\end{aligned}$$

also if $z' = sz$,

$$\begin{aligned}(rz')^\beta = [r(sz)]^\beta = [(rs)z]^\beta = (rs)z^\beta &= (rs)y \\ &= r(sy) = r(sz^\beta) = r(sz)^\beta = r \cdot z'^\beta.\end{aligned}$$

Hence $Z^\beta = Y$: $Z \sim Y$.

From the theory of finitely generated groups it is known that

2.2. *If $(r) \neq 0$ is the least common multiple of the orders of the elements of an abelian group A, then A contains an element of order (r).*

A *primary* group of characteristic a prime ideal (p) is a group in which the order ideal of every element in A is a power of (p). In notation: if x is any element in A, then $O(x) = (p)^{n(x)}$ for some integer $n(x) \geq 0$. The integer $n(x)$ is called the *order-exponent* of the element x. The set of all elements x of any group G such that $O(x) = (p)^{n(x)}$ for some integer $n(x) \geq 0$ is called the *primary component* G_p of G of characteristic (p). It is easy to show that the primary components of a group are subgroups of the given group.

The elements $x_1, x_2, \cdots, x_k$ in a group are said to be *independent* [over R] if the relation $\sum_{i=1}^{k} s_i x_i = 0$ implies $s_i x_i = 0$ for $i = 1, 2, \cdots, k$. Otherwise the elements are said to be *dependent*. An infinity of elements T in A are independent if every finite subset of T forms an independent subset.

The group A is said to be the *direct sum* of the groups $A_1, A_2, \cdots, A_\nu, \cdots$: $A = \sum_\nu A_\nu$, if

(a). A is generated by the A_ν,

(b). $A_\nu \cap B_\nu = 0$ where B_ν is generated by the A_μ for $\mu \neq \nu$. Condition (b) is equivalent to

(b′). Any set of elements $a_\nu \neq 0$ such that a_ν belongs to A_ν is independent.

The following known theorems are listed for future reference.

2. 3. *If A is an abelian group without elements of order 0, then A is a direct sum of its primary components A_p.*

2. 4. *Let $F(A)$ denote the set of elements x in the abelian group A such that $O(x) \neq 0$. Then $F(A)$ is a subgroup of A and $A/F(A)$ contains no element $\neq 0$ of order different from 0.*

If A is a primary abelian group of characteristic (p), denote by $\bar{A}$ the totality of elements x in A such that $px = 0$. The set $\bar{A}$ clearly is a subgroup of A. Consider now a group G having the property that no cyclic subgroup of G has a proper subgroup, i. e., cyclic subgroups are simple. Obviously $\bar{A}$ is such a subgroup. Groups having a field as operator-system are also of this type, and these two are actually the only possibilities.

2. 5. *Let G be a group whose cyclic subgroups are simple, and assume that S and T are subgroups of G such that $S \leq T$. Then S is a direct summand of T.*

2. 6. *If G is a group whose cyclic subgroups are simple, then any greatest independent subset of any subgroup S of G forms a basis of S.*

2. 7. *If G is a group all of whose cyclic subgroups are simple, then any two bases of G contain the same number of elements. [This number may be termed the rank of G].*[9]

2. 8. *If A is a direct sum of cyclic groups of order 0, then any two decompositions of A into direct summands contains the same number of summands.*

2. 9. *Let A be an abelian group and suppose that B is a subgroup of A with the property $rB = B$ for every $r \neq 0$ in R. Then B is a direct summand of A.*[10]

---

[9] Cf. S. MacLane, "A lattice formulation for the transcendence degree and p-bases," *Duke Mathematical Journal*, vol. 4 (1938), pp. 455-468.

[10] Cf. R. Baer, "Abelian groups that are direct summands of every containing abelian group," *Bulletin of the American Mathematical Society*, vol. 46 (1940), pp. 800-806.

In particular it follows that if A is a primary group of characteristic (p) and $r = p$, then B is a direct summand of A.

The primary group $A = \{x_0, x_1, \cdots, x_i \cdots\}$ of characteristic (p) is said to be of type p^∞ if the relation $px_i = x_{i-1}$ holds for every $i > 0$ and if $O(x_0) = (p)$.

2.10. THEOREM.[11] *Let A be a primary group of characteristic (p) with the property $pA = A$. Then A is a direct sum of groups of type p^∞.*

Proof. Let B be a greatest independent subset of $\bar{A}$. By 2.6 B is a basis of $\bar{A}$. Since $pA = A$, there always exists a solution of an equation of the form $px = a$ for any element a in A. If b_ν is any element in B, let $b_{\nu,1}$ be a solution of the equation $px = b_\nu$. Likewise let $b_{\nu,2}$ be a solution of $px = b_{\nu,1}$; and in general if $b_{\nu,j}$ has been defined, let $b_{\nu,j+1}$ be a solution of $px = b_{\nu,j}$, $j = 1, 2, \cdots$. Then $B_\nu = \{b_{\nu,1}, b_{\nu,2}, \cdots, b_{\nu,j}, \cdots\}$ is a subgroup of type p^∞. Let H denote the subgroup of A generated by all the B_ν. If y is an element in H, then y is a sum of elements from a finite number of the B_ν. That the B_ν are independent is easily verified, and hence H is a direct sum of the B_ν. Now H has the property $pH = H$, and by 2.9 it follows that H is a direct summand of A: $A = H + H'$ for a suitable H'. But $\bar{A} \leq H$ and $H \frown H' = 0$, so that $\bar{A} \frown H' = 0$. This implies $\bar{H}' = 0$ and hence $H' = 0$. Therefore $A = H$ and the proof is completed.

3. Quasi-basis.[12] The element x of the primary group A of characteristic (p) is said to have *height* n (where n is an integer ≥ 0) if there exists an element y in A such that $p^n y = x$ and no element z exists in A which satisfies the relation $p^{n+1} z = x$. If every equation of the form $p^i w = x$ has a solution w in A for every $i \geq 0$, the element x is said to have *infinite height* in A. Denote the height of x by $h(x)$. It can be verified that this definition of height does not depend of the generator p of (p).

If A is a primary group, denote by A_i the set of elements x in $\bar{A}$ such that $h(x) \geq i$. It is easily shown that the set of elements A_i is a subgroup of A. Suppose Q is any subset of A. Denote by Q' the set of elements $v' = p^{n(v)-1} v$ for every v in Q where $O(v) = (p)^{n(v)}$. Let Q'_i be that subset of Q' consisting of elements v' for which $n(v) - 1 = i$, i. e., the elements of

---

[11] Cf. L. Zippin, "Countable Torsion Groups," *Annals of Mathematics*, vol. 36 (1935), pp. 86-99; R. Baer, "The subgroups of the elements of finite order of an abelian group," *Annals of Mathematics*, vol. 37 (1936), pp. 766-781. See also [9].

[12] The concept of quasi-basis as used here is essentially a generalization of ideas used by H. Prufer. Prufer also introduced originally the concept of height and groups of type $p\infty$.

Q'_i are of height $\geq i$. The set Q is said to form a *quasi-basis* of A if the following statements hold:

(a). The elements in Q'_i are of exact height i.
(b). Q'_i is a basis of Q_i where $Q_i = \{Q'_i\}$.
(c). $A_i = Q_i + A_{i+1}$ [Direct sum].

It is not difficult to verify that

3.1. *The elements in a quasi-basis Q of A are independent.*

3.2. *If Q is a quasi-basis of A, then it is a basis of A mod pA.*

Proof. (1). The Q_i are independent mod pA.

Assume $\sum_{i=1}^{k} r_i q_i \equiv 0$, $r_i q_i \not\equiv 0 \bmod pA$, r_i in R and for suitable q_i in Q. Then there is a maximum order-exponent m of the q_i in this summation. The above congruence may be written in the form $\sum_{i=1}^{k} r_i q_i = pw$ for some w in A. Now $p^{m-1} r_i = \bar{r}_i p^{m-1}$ for suitable $\bar{r}_i$ in R. Then ${\sum_{i=1}^{k}}' \bar{r}_i (p^{m-1} q_i) = p^m w$, where $\sum'$ means that only terms different from 0 are included in the summation. Now it is clear that the $p^{m-1} q_i$ in this revised summation are in Q'_{m-1} since these q_i are of order (p^m). Hence $\sum_{i=1}^{k} \bar{r}_i p^{m-1} q_i$ is in Q_{m-1} since by condition (c) of the definition of quasi-basis $Q_{m-1} \cap A_m = 0$; and since $p^m w$ is in A_m, it follows that $\sum_{i=1}^{k} \bar{r}_i p^{m-1} q_i = 0$. Furthermore, since the elements $p^{m-1} q_i$ occurring in this summation are elements in Q'_{m-1}, they are independent by (b). Hence every $\bar{r}_i p^{m-1} q_i = 0$. But since $p^{m-1} q_i$ is of order (p), this implies that p divides $\bar{r}_i$ or that $\bar{r}_i \equiv 0 \bmod (p)$. Now $p^{m-1} r_i = \bar{r}_i p^{m-1}$ which is an element in (p^m), so that p divides r_i and $r_i q_i \equiv 0 \bmod pA$. This contradiction implies (1).

(2). Every element in A is mod pA expressible in terms of the q_i.

The proof is given by an induction with regard to the order-exponents of the elements of A. Suppose $O(x) = (p)^{n(x)}$. If $n(x) = 0$, then $x = 0$ and the statement is trivial. Assume that the statement is true for $n - 1$, and that $n(x) = n$. Then $p^{n-1} x$ is in A_{n-1}, and in particular is expressible, mod A_n, in terms of the elements in Q'_{n-1} since $Q_{n-1} + A_n = A_{n-1}$. Hence $p^{n-1} x \equiv \sum_{i=1}^{m} r_i p^{n-1} q_i \bmod A_n$ for suitable $p^{n-1} q_i$ in Q'_{n-1} and r_i in R, $p^{n-1} x = \sum_{i=1}^{m} p^{n-1} \bar{r}_i q_i + y$ where $y = p^n z$ is in A_n and for suitable $\bar{r}_i$ in R. Then $p^{n-1}(x - \sum_{i=1}^{m} r_i q_i$

$-pz) = 0$, i. e., $x - \sum_{i=1} r_i q_i - pz$ is of order-exponent $\leq n-1$, and by the induction hypothesis $x - \sum_{i=1}^{m} r_i q_i - pz \equiv \sum_{j=1}^{m'} s_j g_j$, mod pA, for suitable g_j in Q. This means $x \equiv \sum_{i=1}^{m} r_i q_i + \sum_{j=1}^{m'} s_j g_j$, mod pA, since pz is in pA, and this completes the proof.

3. 3. THEOREM. *The set S is a subset of a quasi-basis of A if, and only if,*

(a). *S is independent.*

(b). $p^i\{S\} = \{S\} \cap p^i A.$

Proof. Let S be any subset of any quasi-basis of A. By 3. 1 the set S is independent. Certainly $p^i\{S\} \leq \{S\} \cap p^i A$. Now any element $x \neq 0$ in $\{S\} \cap p^i A$ has the form $0 \neq x = \sum_{j=1}^{k} r_j s_j = p^i w$ for s_j in S, w in A. It must be proved that x is a p^i-fold of some element in $\{S\}$. Then it may be assumed without loss of generality that none of the (r_j) is divisible by (p^i) since x is in $p^i\{S\}$ if, and only if, $x + p^i y$ for y in $\{S\}$ is in $p^i\{S\}$. Choose $h+1$ as the maximum of the order-exponents of the $r_j s_j$. Now for suitable $\bar{r}_j$ in R it follows that $p^h r_j = r'_j p^h = \bar{r}_j p^{h_j}$ where $\bar{r}_j \equiv 0 \bmod (p)$ and that $h \leq h_j < h + i$. Not every $p^h r_j s_j = \bar{r}_j p^{h_j} s_j$ is 0; and since the s_j belong to a quasi-basis, they are independent by 3. 1. Hence $0 \neq p^{h+i} w = \sum_{j=1}^{k} p^h r_j s_j = \sum_{j=1}^{k} \bar{r}_j p^{h_j} s_j$. If $p^h r_j s_j = 0$, then $p^{h_j} s_j$ is in Q'_{h_j}, and $\sum_{j=1}^{k} \bar{r}_j p^{h_j} s_j$ is an element in A_{h+i}. Since $A = \sum_{m=0}^{h+i-1} Q_m = A_{h+i}$, condition (c) implies that $\sum_{j=1}^{k} \bar{r}_j p^{h_j} s_j = 0$. But the s_j are independent. Consequently $\bar{r}_j p^{h_j} s_j = 0$ and therefore $\bar{r}_j \equiv 0 \bmod (p)$. Hence x is in $p^i\{S\}$.

Conversely, assume that conditions (a) and (b) are true. Let $O(s) = (p)^{n(s)}$ for s in S. Since S is an independent set, $p^{n(s)-1}s$ is of height $n(s)-1$ in $\{S\}$, and hence by (b) it follows that it is of height $n(s)-1$ in A. If S_i is the set of all the s in S such that $n(s)-1=i$, then it follows from (a) that $S_i \cap A_{i+1} = 0$ and that the set S_i is independent. Hence S_i is a part of a basis T_i of $A_i \bmod A_{i+1}$. Denote by D_i the set of all solutions of equations $p^i x = t$ for t in T_i and not in S_i, and denote by Q the set composed of S and D_i. Then Q is a quasi-basis of A containing S as a subset, and this completes the proof.

Remark. This implies the existence of a quasi-basis for every primary group A, since the vacuous set certainly satisfies (a) and (b).

3.4. THEOREM.[18] *If G is a subset of a quasi-basis of A, and if the order-exponents of the elements of G are bounded, then the subgroup generated by G is a direct summand of A.*

Proof. By hypothesis there is an integer k such that $p^k g = 0$ for every g in G. If G is a subset of the quasi-basis Q, denote by T the set of elements in Q not contained in G and consider $A/\{T\}$. Let B be the uniquely determined greatest subgroup of $A/\{T\}$ with the property $pB = B$. Denote by H that subgroup of A for which $H/\{T\} = B$. The proof follows in two parts.

(1). $\{G\} \cap H = 0$.

For assume, that t is an element different from 0 in $\{G\} \cap H$. Since t is in $\{G\}$, t is not in $\{T\}$ for the elements in Q are independent by 3.1. Then $\{T\} + t$ is an element different from 0 in $H/\{T\}$. But $p(H/\{T\}) = H/\{T\}$, and hence t is an element in H and $\{T\} + t \neq 0$ is an element in a subgroup of type p^∞ of $H/\{T\}$. This implies the following relations:

$$\left.\begin{array}{l} t \equiv pw_1 \bmod \{T\} \text{ for some } w_1 \text{ in } H, \\ w_1 \equiv pw_2 \bmod \{T\} \text{ for some } w_2 \text{ in } H, \\ \cdots\cdots\cdots\cdots \\ w_i \equiv pw_{i+1} \bmod \{T\} \text{ for some } w_{i+1} \text{ in } H \\ \cdots\cdots\cdots\cdots \end{array}\right\} \text{ or } \left\{\begin{array}{l} t = pw_1 + t_1 \text{ for } t_1 \text{ in } \{T\}, \\ w_1 = pw_2 + t_2 \text{ for } t_2 \text{ in } \{T\}, \\ \cdots\cdots\cdots\cdots \\ w_i = pw_{i+1} + t_{i+1} \text{ for } t_{i+1} \text{ in } \{T\}, \\ \cdots\cdots\cdots\cdots \end{array}\right.$$

From this the following relation may be derived: $t = t_1 + pt_2 + \cdots + p^{k-1}t_k + p^k w_k$ where k is as chosen above, i. e., $p^k x = 0$ for every x in G. Then $t - \sum_{i=1}^{k} p^{i-1}t_i = p^k w_k$ where the left member is in $\{Q\}$. Therefore by 3.3 there exists a z in $\{Q\}$ such that $p^k z = t - \sum_{i=1}^{k} p^{i-1}t_i$. Since $p^k\{G\} = 0$, $p^k z$ is in $\{T\}$ and this implies that t is in $\{T\}$. But t is also in $\{G\}$ and $\{G\} \cap \{T\} = 0$. Hence $t = 0$, and (1) is true.

(2). Every element in A is expressible as the sum of an element in $\{G\}$ and an element in H.

For let v be any element in A. From 3.2 it follows that $v \equiv \sum_{i=1}^{h} r_i q_i$, mod pA, for suitable q_i in Q and r_i in R. Then $v = v_1 + pw$ for $v_1 = \sum_{i=1}^{h} r_i q_i$ and w in A. Then clearly $v \equiv pw \bmod \{Q\}$. This means that every element in A is a p-fold mod $\{Q\}$, or that $A/\{Q\} = p(A/\{Q\})$. [Henceforth for purposes of notation in this section S^* will denote the quotient group $S/\{T\}$ for any subgroup S between $\{T\}$ and A.] Since $A^*/\{Q\}^* = A/\{Q\}$, this implies $A^*/\{Q\}^* = p(A^*/\{Q\}^*)$. Let v^* be any element in A^*. Then from the above statement it follows that v^* belongs to a subgroup of A^* of type p^∞ mod $\{Q\}^*$ unless $A^* = \{Q\}^*$ in which case there is nothing to prove. From this one gets the following relations:

---

[18] Cf. R. Baer, "Primary abelian groups and their automorphisms," *American Journal of Mathematics*, vol. 59 (1937), pp. 100-101.

$$\left.\begin{aligned} v^* &\equiv pw^*_1 \bmod \{Q\}^* \text{ for } w^*_1 \text{ in } A^*,\\ w^*_1 &\equiv pw^*_2 \bmod \{Q\}^* \text{ for } w^*_2 \text{ in } A^*,\\ &\cdots\cdots\cdots\cdots\cdots\\ w^*_{i-1} &\equiv pw^*_i \bmod \{Q\}^* \text{ for } w^*_i \text{ in } A^*,\\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}\right\} \text{ or } \left\{\begin{aligned} v^* &= pw^*_1 + t^* \text{ for some } t^* \text{ in } \{Q\}^*,\\ w^*_1 &= pw^*_2 + t^*_1 \text{ for some } t^*_1 \text{ in } \{Q\}^*,\\ &\cdots\cdots\cdots\cdots\cdots\\ w^*_{i-1} &= pw^*_i + t^*_{i-1} \text{ for some } t^*_{i-1} \text{ in } \{Q\\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}\right.$$

From these relations one can derive: $v^* = t^* + pt^*_1 + p^2t^*_2 + \cdots + p^{k-1}t^*_{k-1} + p^kw^*_k$ where k is as chosen above: $p^k\{Q\}^* = 0$. That $h^* = p^kw^*_k$ is an element in H^* is shown as follows. Remember that $\{Q\} = \{G\} + \{T\}$ and hence that $\{Q\}^* \sim \{G\}$, so that $p^k\{Q\}^* = 0$. Then since t^*_j is an element in $\{Q\}^*$, by repeated application of the above process it follows that

$$\begin{aligned} h^* &= p^k(pw^*_{k+1} + t^*_k) = p(p^kw^*_{k+1}),\\ p^kw^*_{k+1} &= p^k(pw^*_{k+2} + t^*_{k+1}) = p(p^kw^*_{k+2}),\\ &\cdots\cdots\cdots\cdots\cdots\\ p^kw^*_{k+i} &= p^k(pw^*_{k+i+1} + t^*_{k+i}) = p(p^kw^*_{k+i+1}),\\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

This shows that h^* is an element of a group of type p^∞, and since H^* is the greatest subgroup of A^* such that $H^* = pH^*$, it follows that h^* is contained in H^*. As v^* was chosen as any element in A^* and since $m^* = t^* + pt^*_1 + \cdots + p^{k-1}t^*_{k-1}$ is an element in $\{Q\}^*$, the fact that h^* is in H^* shows that $A^* = \{H^*, \{Q\}^*\}$. Then if v is any element in A, it follows that $v + \{T\} = h + \{T\} + g + \{T\}$ for some h in H and g in $\{G\}$. Then $v = h + \bar{t}_1 + g + \bar{t}_2$ for suitable $\bar{t}_i$ in $\{T\}$. Or $v = \bar{h} + g$ where $\bar{h} = h + \bar{t}_1 + \bar{t}_2$. Since $\{T\} \leq H$, $\bar{h}$ is in H and since g is in $\{G\}$, this shows $A = \{H, G\}$. This completes the proof of the theorem.

The above remarks and arguments also imply the following

3.5. COROLLARY. *If Q is a quasi-basis of A, if G is a subset of Q such that the order-exponents of the elements in G are bounded, if T is the complement of G in Q, and if H is the greatest subgroup of A such that $\{T\} \leq H$ and $p(H/\{T\}) = H/\{T\}$, then $A = H + \{G\}$.*

3.6. Corollary. *If A possesses a quasi-basis Q such that the order-exponents of its elements are bounded, then $A = M + H$ where H is a direct sum of groups of type p^∞ and M is a direct sum of cyclic groups of bounded order-exponents.*

3.7. Corollary. *If the quasi-bases of A are finite and if A contains no element $\neq 0$ of infinite height, then A is a direct sum of a finite number of cyclic groups.*

3.8. Corollary. *If the order-exponents of the elements of A are bounded, then A is a direct sum of cyclic groups.*

In the remainder of this section statements are given which are known to hold in the classical theory. In most cases the generalizations are in the literature or are trivial, but the statements are given since they are used in subsequent proofs.

3.9. *Let $A \neq 0$ be a primary group of characteristic (p) and suppose that A does not contain a cyclic subgroup $[\neq 0]$ which is itself a direct summand of A. Then A is a direct sum of groups of type p^∞.*

3.10. *The primary group A possesses a cyclic quotient group if, and only if, A possesses a cyclic direct summand.*

3.11. Theorem. *If A is a primary group of characteristic (p), then A is homomorphic to a group of type p^∞ if, and only if, the order-exponents of the elements in A are unbounded.*[14]

3.12. *Let A be a group such that any two of its elements are dependent and which contain only elements of order 0. Then A is cyclic if, and only if, A is generated by a finite number of elements.*

3.13. Theorem. *If A is generated by a finite number of elements, then every subgroup and every quotient group of A is a direct sum of a finite number of cyclic groups.*[15]

4. Character groups. Assume, as before, that R is a ring satisfying conditions (A) and (B). Consider another ring R' such that R and R' are anti-isomorphic, i. e., a). There is a one-to-one correspondence [denoted by $'$] between the elements r in R and r' in R' such that:

---

[14] For a discussion of this theorem in the classical case see R. Baer, "Dualisms in abelian groups," *Bulletin of the American Mathematical Society*, vol. 43 (1937), pp. 121-124.

[15] Cf. the reference in footnote 7. See also G. Kothe, "Die Struktur der Ringen deren Restklassenring nach dem Radical vollstandig ist," *Math. Zeitschrift*, vol. 32 (1930), pp. 161-186.

b). $(r+s)' = r' + s'$ for every r, s in R,

c). $(rs)' = s'r'$ for every r, s in R.

Let A and V be two abelian groups admitting the ring R as left operators. Moreover, let V also admit R' as a system of right operators so that V and the rings R and R' are subject to the following conditions:

d). $(u+v)r' = ur' + vr'$ for u, v in V and r' in R',

e). $(vr')s' = v(r's')$ for v in V and r', s' in R',

f). $s(vt') = (sv)t'$ for s in R, v in V, and t' in R'.

A further hypothesis imposed on the rings R, R' is the following:

(K) 1. *If v is any element in V and (r) is a left ideal in R, then there exists a right ideal (s') in R' [depending on (r) and v] satisfying:*

i. $(r)v = v(s')$; ii. $(s) \leq (r)$.

2. *If v is any element in V and (r') is any right ideal in R', then there exists a left ideal (s) in R [depending on (r') and v] satisfying:*

i. $v(r') = (s)v$; ii. $(s') \leq (r')$.

Under this set-up it is easy to see that a subgroup U of V is left-cyclic if, and only if, it is right-cyclic. Moreover, the same generator may be used for both cases. Every left-subgroup of V is also a right-subgroup and conversely. The following notation will be used:

$O(x)$ = left-order of an element x in V or in A.

$O_r(v)$ = right-order of an element v in V. That is $O_r(v)$ is the set of all elements s' in R' such that $vs' = 0$.

(r, B) = totality of elements x in B such that $rx = 0$ for a fixed r in R.

(B, r') = totality of elements v in $B \leq V$ such that $vr' = 0$ for fixed r' in R'.

It easily follows that if r is an element in R, then $(r, V) = (V, r')$ is a two-sided subgroup of V. And it is a consequence of condition (K) that if v is an element in V and t an element [0 or not] in R, then t is in $O(v)$ if, and only if, t' is in $O_r(v)$.

The mapping f of A into V is termed a *character* of A in V if the following conditions are satisfied:[4]

1). $f(x) = v$ where v is a uniquely determined element in V for every x in A,

2). $f(x+y) = f(x) + f(y)$ for every x, y in A,

3). $f(rx) = f(x) \cdot r'$ for every r in R, x in A.

From these conditions it is clear that f is a linear mapping of A into a subgroup of V. This definition of character coincides with the classical definition if R and R' are replaced by the ring of ordinary integers. Then $R' = R$ and $nv = vn$ for every v in V and every integer n, in which case condition (K) holds. In order that the set of characters of A in V may form an additive group also admitting the ring R as left operators, the further definitions are given:

4). If f and g are characters of A in V, put $(f+g)(x) = f(x) + g(x)$ for every x in A,

5). $(rf)(x) = r \cdot f(x)$ for every r in R, x in A, and every character f of A in V. The group of all characters of A in V thus defined will be denoted by $C_V(A)$.

Let A be an abelian group and suppose that A' is a subgroup of A. If f is a character of A in V and f' a character of A' in V such that $f(x) = f'(x)$ for every x in A', then the character f of A in V is said to *induce* the character f' of A' in V. Every character of A in V induces a character of every subgroup of A in V. Conversely it is not difficult to derive the following:

4. 1. *Let* $A = \sum_{\nu} A_{\nu}$ *and let* f_{ν} *be a character of* A_{ν} *in* V. *Then there exists one, and only one, character* f *of* A *in* V *such that* $f(x) = f_{\nu}(x)$ *for every* x *in* A_{ν}.

4. 2. *If* $A = A_1 + A_2 + \cdots + A_n$, *then* $C_V(A) = C_V(A_1) + C_V(A_2) + \cdots + C_V(A_n)$. *Here the* $C_V(A_i)$ *have been identified with the groups* K_i *of all those characters of* A *in* V *which map every element in* $\sum_{j \neq i} A_j$ *on* 0.

4. 3. *If* $V = V_1 + V_2 + \cdots + V_n$, *then* $C_V(A) = C_{V_1}(A) + C_{V_2}(A) + \cdots + C_{V_n}(A)$.

Remark. It is to be noted that these last two theorems apply only to cases in which only a finite number of summands is involved. No statement is made concerning the infinite number of summands which is actually different.

4. 4. *Every character of* A/B *in* V *for any subgroup* B *of* A *is induced by one, and only one, character of* A *in* V *which maps* B *on* 0.

Proof. If f is a character of A/B in V, define $g(x) = f(B + x)$ for every x in A. [If x is in B, $g(x) = 0$.] Then g is the desired character of A in V.

5. Sufficient conditions.

5.1. LEMMA.[16] *If A is a cyclic group of order (r), where (r) is* 0 *or not, then $C_V(A) \sim (V, r') = (r, V)$.*

Proof. Let $A = \{z\}$, so that $O(z) = (r)$. If f is any character of A in V, $f(z) = v$ where v is an element in V. Then $0 = f(0) = f(rz) = f(z) \cdot r' = vr'$, and this proves that v is contained in (V, r'). Let u be an arbitrary element in (V, r') and define $g(sz) = us'$ for every element s in R. [In particular $g(z) = u$.] Then g is a character of A in V, for:

a). $sz = tz$ if, and only if, $s - t \equiv 0 \bmod (r)$. But this implies $s' - t' \equiv 0 \bmod (r')$ and hence $u(s' - t') = 0$, which gives $us' = ut'$. This shows that g is a single-valued function.

b). If $x = sz$, $y = tz$, then $g(x + y) = g(sz + tz) = g[(s + t)z] = u(s + t)' = u(s' + t') = g(sz) + g(tz) = g(x) + g(y)$.

c). $g(rx) = g(rsz) = u(rs)' = u(s'r') = (us')r' = g(sz)r' = g(x) \cdot r'$.

These three properties show that g is a character of A in V. Hence there is a $(1, 1)$-correspondence between the elements of $C_V(A)$ and (V, r') obtained by mapping each character f in $C_V(A)$ upon the element $f(z)$ in (V, r'). This correspondence defines an isomorphism; hence $C_V(A) \sim (V, r')$, as was to be proved.

5.2. COROLLARY. *If A is cyclic of order (r), then $C_V(A) \sim A$, if, and only if, (V, r') $[= (r, V)]$ is cyclic of order (r).*

5.3. THEOREM. *If A is a direct sum of a finite number of cyclic groups: $A = Z_1 + Z_2 + \cdots + Z_m$, where $O(Z_i) = (r_i)$, and if (V, r') is cyclic of order (r_i), then $C_V(A) \sim A$.*

Proof. By 4.2 it follows that $C_V(A) = C_V(Z_1) + C_V(Z_2) + \cdots + C_V(Z_m)$, and the theorem is a direct consequence of 5.2.

Remark. Since $(0, V) = V$, it follows that either *all* or *none* of the $(r_i) = 0$. For suppose $(r_1) = 0$. Then $(r_1, V) = (V, r'_1) = V$ and hence must be cyclic of order $(r_1) = 0$. If now $(r_i) \neq 0$ for any $i > 1$, then it is impossible that (V, r'_1) be cyclic of order (r_i).

---

[16] For a treatment of the classical case of this theorem see the reference by P. Alexandroff and H. Hopf mentioned in footnote 3.

The following is an example which shows that A may contain elements of order 0 and also elements of order different from 0 and still have the isomorphism condition hold: Let R be the ring of ordinary integers, M the additive group of rational numbers, and Z a cyclic group of order a prime number p. Take $A = V = M + Z$. Then $C_V(A) \sim A$, and A contains elements of order 0 and also elements of order $p \neq 0$. However, it is to be noted that A is not a direct sum of cyclic groups.

6. Necessary conditions.

6.1. LEMMA. *Let A be an abelian group containing an element of order $(r) \neq 0$. Then the condition $C_V(A) \sim A$ implies that V contains an element of order (r).*

Proof. Since $C_V(A) \sim A$, there exists a character f of A in V of order (r). If $O[f(x)] = (m_x)$, then $(m_x) \geqq (r)$ since $r \cdot f(x) = 0$. Let m be the least common multiple of the (m_x); then $(m) \geqq (r)$. Now $m \cdot f(x) = 0$ for every x in A, and therefore $(r) \geqq (m)$. These results imply $(m) = (r)$. Define V' as the group generated by all the $f(x)$ as x runs through A. Then V' is abelian with the least common multiple of its order-ideals $= (r) \neq 0$. Hence, by 2.2, V', and therefore V, contains an element of order (r) as was to be proved.

6.2. THEOREM. *Let R be a field and suppose that $V = Rv = vR' \neq 0$ is cyclic. Then if $B = \sum_{\nu} b_\nu$ is a direct sum of $\aleph$ cyclic groups different from 0, $C_V(B)$ contains $2^{\aleph}$ independent characters.*

Proof. Let $P(R)$ denote the prime field contained in R. Define $\bar{B}$ as the totality of sums of the form $\sum_{i=1}^{n} p_i b_{\nu_i}$ for p_i in $P(R)$. Put $\bar{V} = P(R)v$, and note that $\bar{B}$ and $\bar{V}$ are not necessarily subgroups. Then a character of $\bar{B}$ in $\bar{V}$ may be defined by mapping b_ν either on 0 or on v. Clearly $2^{\aleph}$ such characters may be defined since there are $\aleph$ of the $b_\nu \neq 0$, and since the b_ν form a basis of $\bar{B}$. Now $P(R)$ contains at most $\aleph_0$ elements, since it is a prime field, and $C_{\bar{V}}(\bar{B})$ admits the countable field $P(R)$ as operators. Hence the number of elements in a basis of $C_{\bar{V}}(\bar{B})$ is equal to the number of elements in $C_{\bar{V}}(\bar{B})$ if this number is infinite. This shows there are $2^{\aleph}$ independent characters in $C_{\bar{V}}(\bar{B})$. Note that every character of $\bar{B}$ in $\bar{V}$ is induced by one, and only one, character of B in V. Assume that the characters $g_1, g_2, \cdots, g_n$ of B in V induce the characters $\bar{g}_1, \bar{g}_2, \cdots, \bar{g}_n$ of $\bar{B}$ in $\bar{V}$. It can now be proved without difficulty that the characters $\bar{g}_1, \bar{g}_2, \cdots, \bar{g}_n$ are independent if, and only if, the characters $g_1, g_2, \cdots, g_n$ are independent. Hence since the group $C_{\bar{V}}(\bar{B})$ contains $2^{\aleph}$ independent characters, so does the group $C_V(B)$. This completes the proof.

6. 3. Corollary. *Let A be a primary group of characteristic (p). If $A = \sum_\nu A_\nu$ where there are $\aleph$ summands each of which is cyclic and different from 0, and if V contains an element of order (p), then there are $2^{\aleph}$ independent characters of A in V.*

Proof. Note that $A/pA = \sum_\nu A_\nu / \sum_\nu pA_\nu = \sum_\nu (A_\nu/pA_\nu) = \sum_\nu B_\nu$; and since every A_ν/pA_ν is cyclic of order (p), every B_ν is cyclic and of order (p). Then one may as well consider $R/(p)$ [which is a field by 1. 4] instead of R as the left-operator system of A/pA and of $\bar{V} = Rv \leq V$ where v is an element in V of order (p). For since $\bar{V}$ is left-cyclic of order (p), by Section 4 it is of right-order (p'). Hence $\bar{V}$ admits $[R/(p)]' = R'/(p')$ as a system of right-operators.

By 4. 4 every character of A/pA in $\bar{V}$ is induced by a character of A in $\bar{V}$ which maps pA on 0. Conversely every character of A in $\bar{V}$ which maps pA on 0 induces one, and only one, character of A/pA in $\bar{V}$. Hence a set of characters of A which map pA on 0 is independent if the corresponding set of characters of A/pA is independent. It follows by 6. 2 that there exist $2^{\aleph}$ independent characters of A/pA in $\bar{V}$. Hence there are $2^{\aleph}$ independent characters of A in $\bar{V}$ [and therefore in V since $\bar{V} \leq V$].

6. 4. Corollary. *Let $A = \sum_\nu A_\nu$ where A_ν is cyclic of order 0 and where there are $\aleph$ summands. If $V \neq 0$, $C_V(A)$ contains $2^{\aleph}$ independent characters.*

Proof. Case 1. V contains an element of order a prime ideal (p).

A/pA contains only elements of order (p). Hence by 2. 6 A/pA is a direct sum of $\aleph$ cyclic groups of order (p). Then by 6. 3 there are $2^{\aleph}$ independent characters of A/pA in V. The statement for Case 1 is now a consequence of 4. 4.

Case 2. V contains an element w of order 0.

If $W = \{w\}$, then $C_V(A) \geq C_W(A)$. If R is a field, the statement follows from 6. 2. If R is not a field, then R contains a prime ideal (p) and $R/(p)$ is a field. Now $C_{W/pW}(A)$ consists only of characters f [except 0] of order (p); hence if $f(x) = v \neq 0$ for some element x in $R/(p)$, then $O(v) = (p)$. By Section 4 $O_r(v) = (p')$; and since $f(py) = f(y) \cdot p'$, it follows that $C_{W/pW}(A/pA) = C_{W/pW}(A)$. But W/pW is of order (p) and Case 1 implies that there are $2^{\aleph}$ independent characters in $C_{W/pW}(A/pA) = C_{W/pW}(A)$.

Choose a representative $r(x)$ in every class x of $R/(p)$. Let $\bar{g}$ be any character in the above set of $2^{\aleph}$ independent characters of A in W/pW. If the elements a_ν, for a_ν in A_ν, form a basis of A and if $w^* = W/pW$, let $\bar{g}(a_\nu) = w^* s'$ for some s' in $R'/(p')$. Then there exists one, and only one, character g of A in $W = \{w\}$ such that $g(a_\nu) = w \cdot r(s')$ for every a_ν. The proof is now completed by showing that the set of characters g of A in W thus obtained is independent.

6.5. LEMMA. *If $A = \sum_p A_p$ where the A_p are primary groups and the summation is over distinct prime ideals (p) in R, then $C_V(A_p)$ contains no element of order 0 if, and only if, $C_V(A)_p$ and $C_V(A_p)$ are essentially the same.*

Proof. Let g be a character of A in V of order (p^n) for some integer n, and suppose that x is an element in A_q, $(q) \neq (p)$, of order (q^m). Then (p^n) and (q^m) are relatively prime and there exist elements $\bar{p}$ and $\bar{q}$ in R such that $\bar{p}p^n + \bar{q}q^m = 1$. Hence $g(x) = [\bar{p}p^n + \bar{q}q^m]g(x) = \bar{p}p^n \cdot g(x) + \bar{q}q^m \cdot g(x) = \bar{q} \cdot g(x) \cdot q'^m s' = \bar{q} \cdot g(q^m x) \cdot s' = 0$, using condition (K). This shows that g is the zero character of A_q in V for every $(q) \neq (p)$. Therefore $C_V(A)_p \leq C_V(A_p)$. [Actually $C_V(A)_p \sim G \leq C_V(A_p)$]. On the other hand suppose that f is a character of A_p in V of order different from 0. Let $O(f) = (r) = (p^\alpha q)$ where (p) and $(q) \neq R$ are relatively prime and $\alpha \geq 0$. Clearly $O(p^\alpha f) = (q)$. Hence there is an element x in A_p such that $(p^\alpha f)(x) = v \neq 0$ for some v in V. If $O(x) = (p)^n$, then $p^n v = v p'^n s' = (p^\alpha f)(x) \cdot p'^n s' = p^\alpha \cdot f(p^n x) \cdot s' = p^\alpha \cdot f(0) \cdot s' = 0$; hence $O(v) \geq (p^n)$. But also $q \cdot v = q(p^\alpha f)(x) = (qp^\alpha) \cdot f(x) = 0$, and hence $O(v) \geq (q)$. Since (q) and (p^n) are relatively prime, this implies $O(v) = R$ or that $v = 0$. This contradiction shows the assumption that $O(f) = (r)$ where $(r) = (p^\alpha q)$ for $q \neq R$ and $(p) \cap (q) = 0$ is false. Hence $(q) = R$ and $O(f)$ is a power of the prime ideal (p). This proves that f belongs to $C_V(A)_p$, so that $C_V(A_p) \leq C_V(A)_p$. Hence $C_V(A_p) = C_V(A)_p$, and the proof is complete.

6.6. THEOREM. *Let $A = \sum_p A_p$ where the A_p are primary groups and the summation is over distinct prime ideals in R. Then the condition $C_V(A) \sim A$ implies*

a). *$C_V(A_p) \sim A_p$ for every (p),*
b). *Almost every $A_p = 0$.*

Proof. a). Let (p) be a fixed prime ideal in R. Then $A = A_p + B_p$ where $B_p = \sum_{p \neq q} A_q$, and 4. 2 implies that $C_V(A) = C_V(A_p) + C_V(B_p)$. Since there are no characters of order 0 in $C_V(A)$ [because of the condition of isomorphism of A and its character group], it follows by 6. 5 that $C_V(A)_p = C_V(A_p)$. But since $A_p \sim C_V(A)_p$, this shows that $A_p \sim C_V(A_p)$ as was to be proved. Note that if $A_p \neq 0$, a). implies that $C_V(A_p) \neq 0$.

b). Assume that there exists an infinity of distinct prime ideals $(p_1), (p_2), \cdots$ such that $A_{p_i} \neq 0$. Then by a). $C_V(A_{p_i}) \neq 0$; and hence there exists a character $f_{p_i} \neq 0$ in $C_V(A_{p_i})$. By 4. 1 there is a character f of A in V such that $f(x) = f_{p_i}(x)$ for every x in A_{p_i}. Suppose $O(f) = (r) \neq 0$. Then $(r) \leq O(f_{p_i}) \leq (p_i)$ for every i. But since $(r) \neq 0$ is the product of only a finite number of prime ideals by 1. 5, this is impossible. Hence $(r) = 0$, and then it is impossible that $C_V(A) \sim A$ since A contains no element of order 0. This contradiction implies that only a finite number of the A_{p_i} are different from 0.

6. 7. Corollary. *Suppose $A = \sum_{\nu} A_\nu$ are cyclic and either all are of order 0 or all are of order different from 0. Then the condition $C_V(A) \sim A$ implies that the number of summands different from 0 is finite.*

Proof. Case 1. A contains only elements of order 0.

Assume there are $\aleph$ of the A_ν different from 0. Then V is not vacuous and 6. 4 implies that there are $2^{\aleph}$ independent characters of A in V. Since there are $\aleph$ elements in one basis of A, 2. 8 implies the maximum number of independent elements in A is $\aleph$. Then it is impossible that $C_V(A) \sim A$.

Case 2. A contains no element of order 0.

If A contains an element of order (p), it follows by 6. 1 that V also contains an element of order (p). By 6. 6 there are only a finite number of the primary components $A_p \neq 0$. It is a consequence of 2. 3 that each A_ν is a direct sum of cyclic groups of prime-power orders, and therefore A_p is a direct sum of cyclic groups. Then 6. 3 implies that there are $2^{\aleph}$ independent characters of A_p in V. Hence by 2. 7 it is impossible that $C_V(A_p) \sim A_p$ in contradiction to 6. 6.

6. 8. Lemma. *If A and V are primary groups of left-type p^∞, then $C_V(A) \neq 0$ and contains only characters (except 0) of order 0.*

Proof. Since every left-cyclic subgroup of V is also right-cyclic, as follows from Section 4, it is clear that V is also of right-type p^∞. Then A and V have the forms: $A = \{x_0, x_1, \cdots, x_i, \cdots\}$ and $V = \{v_0, v_1, \cdots, v_i, \cdots\}$

where $O(x_0) = (p)$, $O(v_0) = (p')$, $px_i = x_{i-1}$, and $v_i p' = v_{i-1}$. Any element y in A has the form $y = rx_i$ where r is uniquely determined mod (p^{i+1}) and i is uniquely determined by y. A similar statement follows for the element v. Define $g(sx_i) = v_i s'$ for every s in R $(i = 0, 1, 2, \cdots)$. It is easy to show that g is a character of A in V different from 0.

Let f be any character of A in V different from 0 and assume $f(x_i) = v \neq 0$ for a suitable i and some v in V. Now for every k, $p^k x_{i+k} = x_i$. Hence, by condition (K), $0 \neq v = f(x_i) = f(p^k x_{i+k}) = f(x_{i+k}) \cdot p'^k = \bar{v}_k \cdot p'^k = m_k p^k \bar{v}_k = m_k p^k \cdot f(x_{i+1}) \neq 0$. Then, $O(f)$ does not divide (p^k). Since this holds for every k no matter how large, it follows that $O(f) = 0$. This completes the proof.

6.9. LEMMA. *If A is a primary abelian group of characteristic (p), then $C_V(A)$ cannot contain elements $\neq 0$ of infinite height in $C_V(A)$.*

Proof. Let f be a character of A in V such that there exist characters f_i of A in V with the property $p^i f_i = f$ for every $i > 0$. Suppose that x is an arbitrary element in A. If $O(x) = (p^n)$, then $f(x) = (p^m f_m)(x) = [f_m(x)][p'^m k'_m] = [f_m(p^m x)]k'_m = f_m(0) \cdot k'_m = 0$ for a suitable k'_m in R' by condition (K).

6.10. COROLLARY. *If A is a primary group of characteristic (p), the condition $C_V(A) \sim A$ implies that A does not have elements different from 0 of infinite height in A.*

6.11. THEOREM. *If A is a primary group of characteristic (p) and if $C_V(A) \sim A$, then the order-exponents of the elements of A are bounded.*

Proof. By 6.10 there are no elements in A of infinite height. Assume that the order-exponents of the elements of A are not bounded. Since $C_V(A) \sim A$, it may be assumed without loss of generality that V is a primary abelian group of characteristic (p). These assumptions imply that the order-exponents of the elements of V are also not bounded. For if $(q) = (p^a)$ is an ideal in R such that $qV = 0$, then certainly $q \cdot C_V(A) = 0$, and the condition of isomorphism implies $q \cdot A = 0$ which is contrary to assumption.

Suppose that $V = Z + U$ where Z is cyclic of order (p^n) for some integer $n > 0$. Then 4.3 implies that $C_V(A) = C_Z(A) + C_U(A)$. By 4.4 every character of A/p^nA in Z is induced by a character of A in Z. Furthermore, if f is a character of A in Z of order (p^n), it follows, on using condition (K), that $f(p^n x) = f(x) \cdot p'^n = sp^n \cdot f(x) = 0$ for every x in A. Hence f maps p^nA on the zero-element of Z. These statements prove $C_Z(A) = C_Z(A/p^n{}_A)$.

Similarly $C_{p^{n-1}Z}(A) = C_{p^{n-1}Z}(ApA)$, and $C_{p^{n-1}Z}(A)$ is the subgroup of elements of order (p) in $C_Z(A)$. From the fact that A contains no element different from 0 of infinite height in A and the assumption that the order-exponents of the elements of A are not bounded, it follows that A/pA contains an infinity, say $\aleph$, of independent elements. For if A/pA contained only a finite number of independent elements, then every quasi-basis of A would contain only a finite number of independent elements. And in this case 3. 7 implies that the order-exponents of the elements of A are bounded.

Since by 3. 2 a quasi-basis of A is a basis of A/pA, 6. 3 implies that $C_{p^{n-1}Z}(A/pA) = C_{p^{n-1}Z}(A)$ contains $2^{\aleph}$ independent characters. [Use 6. 1 to show the existence of an element in V of order (p).] Then $C_Z(A)$ contains at least $2^{\aleph}$ independent characters of order (p). Now the conditions $C_V(A) \sim A$ and $C_V(A) = C_Z(A) + C_U(A)$ imply that $A = B + D$ where $B \sim C_Z(A)$ and hence has the order-exponents of its elements bounded and contains $2^{\aleph}$ independent elements. Then $A/pA = B/pB + D/pD$. Now B, and hence pB, contains $2^{\aleph}$ independent elements of order (p), while by assumption A/pA contains only $\aleph$ independent elements. [See 2. 7]. These facts lead to the contradiction $2^{\aleph} \leq \aleph$.

Then the assumption that V contains a cyclic direct summand of order different from 0 is false. This implies [using 3. 9] that V is a direct sum of groups of type p^{∞}. Now since the order-exponents of the elements of A are not bounded, it follows by 3. 11 that there exists a subgroup W of A such that A/W is a group of type p^{∞}. By 6. 8 there is a character g of A in V of order 0. Then because of the isomorphism condition, A also contains an element of order 0. This contradiction completes the proof.

7. Characters whose values form a cyclic group of order 0.

7. 1. THEOREM. *Let V be cyclic of order* 0. *Then there exist n independent characters of A in V if, and only if, there exists a direct summand of A which is itself a direct sum of n cyclic groups of order* 0.

Proof. Suppose that $A = B + D$ where B is a direct sum of n cyclic groups of order 0. By 4. 2 $C_V(A) = C_V(B) + C_V(D)$, and 5. 3 implies that $C_V(B)$ is a direct sum of n cyclic groups of order 0. This proves the sufficiency.

The necessity may be proved by complete induction with regard to n. The statement certainly holds for $n = 0$. Assume that there exist n independent characters of A in V and that the theorem holds for $n - 1$. Then there exist $n - 1$ independent characters of A in V, and by the induction hypothesis $A = B + D$ where D is a direct sum of $n - 1$ cyclic groups of order 0. If every character of A in V would map B on the 0-element of V, the character group of A in V would be the same as the character group of

D in V. The latter is by 5.3 a direct sum of $n-1$ cyclic groups of order 0 since D is the direct sum of $n-1$ cyclic groups of order 0. But this implies that $C_V(A)$ is a direct sum of $n-1$ cyclic groups, and by 2.8 this is impossible since by hypothesis $C_V(A)$ contains n independent elements.

Thus there exists a character f of A in V and an element b in B such that $f(b) \neq 0$. Denote by B' the set of all elements x in B such that $f(x) = 0$. Then B' is a proper subgroup of B (possibly $B' = 0$), and f defines an "anti-isomorphism" of B/B' upon a subgroup of the cyclic group V. As right ideals in R' are principal right ideals, this subgroup V' of V is cyclic of order 0, i. e., $V' = vR'$ for some element v in V. Denote by g some element in B such that $f(g) = v$. That B is a direct sum of B' and the cyclic group generated by g is proved as follows. Assume first that x is any element in B. Then there exists some element r in R such that $f(x) = vr'$. Hence $f(x-rg) = f(x) - f(g) \cdot r' = 0$, i. e., $x - rg$ is an element in B'; and B is generated by g and B'. Suppose, furthermore, that u is an element of B' and that s is an element in R such that $u + sg = 0$, then $0 = f(u+sg) = f(u) + f(sg) = f(g) \cdot s' = vs'$, as u is in B'. Since v is of order 0, this implies that $s = 0$; hence $u = 0$ too. This proves that $A = B' + \{g\} + D$, i. e., A possesses a direct summand $D + \{g\}$ which is a direct sum of n cyclic groups of order 0.

Let B be a group all of whose elements $\neq 0$ are of order 0. B is said to be of finite rank if there exists an integer $n \geq 0$ such that any $n+1$ elements in B are dependent. The smallest integer n of this type is the rank of B: $r(B) = n$.

7.2. COROLLARY. *Suppose that V is cyclic of order 0 and that all the elements of A (other than the 0-element) are of order 0. Then $r(A) = r[C_V(A)]$ is finite if, and only if, A is a direct sum of a finite number of cyclic groups of order 0.*

Proof. Suppose that $r(A) = r[C_V(A)] = n$. Then by 7.1 $A = B + D$ where B is the direct sum of n cyclic groups of order 0. Since $r(A) = r(B)$, it follows that $D = 0$. The converse is a direct consequence of 5.3.

An example may be given to show that the Corollary does not hold if the rank of A is not finite. Let M be isomorphic to the group of rational numbers [taking R as the domain of ordinary integers]. If V is cyclic of order 0, note that $C_V(M) = 0$. Let B be a direct sum of $\aleph$ cyclic groups of order 0. Then by 6.4 $r[C_V(B)] = \aleph'$ where $\aleph' > \aleph$. Put D equal to the direct sum of $\aleph'$ groups isomorphic to M, and define $A = B + D$. Then $C_V(A) = C_V(B) + C_V(D) = C_V(B)$ since $C_V(D) = 0$. But $r[C_V(A)] = r(A)$, and A is certainly not a direct sum of cyclic groups.

8. The main theorems.

8.1. THEOREM. *Let A be an abelian group without elements of order* 0. *Then $C_V(A) \sim A$ if, and only if, the following two conditions hold:*

a). *A is a direct sum of a finite number of cyclic groups.*

b). *If A contains an element of order (r), then the subgroup $(V, r') = (r, V)$ of V is cyclic of order (r).*

Proof. The sufficiency follows from 5.3. Assume that $C_V(A) \sim A$. Then by 2.3 $A = \sum_p A_p$ where A_p is the primary component of A of characteristic (p) and the summation is taken over all distinct prime ideals in R. From 6.6 it follows that $C_V(A_p) \sim A_p$ for every (p) in R. Theorem 6.11 now implies that the order-exponents of the elements of A_p are bounded; 3.8 shows that $A_p = \sum_\nu Z_{p\nu}$ where $Z_{p\nu}$ is cyclic. Then $A = \sum_p A_p = \sum_p \sum_\nu Z_{p\nu}$, which states that A is a direct sum of cyclic groups of orders different from 0. By 6.7 only a finite number of the $Z_{p\nu}$ are different from 0, and this completes the proof of a).

As noted above, A is a direct sum of its primary components; and since A is [by part a) already proved] a direct sum of a finite number of cyclic groups, only a finite number of the primary components A_{p_i} of A are different from 0. Then $A = A_{p_1} + A_{p_2} + \cdots + A_{p_n}$ for some integer n and $A_{p_i} \neq 0$, $(p_i) \neq (p_j)$ for $i \neq j$. By 4.2 it follows that $C_V(A) = C_V(A_{p_1}) + C_V(A_{p_2}) + \cdots + C_V(A_{p_n})$, and 6.6 implies that $C_V(A_{p_i}) \sim A_{p_i}$ for $i = 1, 2, \cdots, n$. Since A is a direct sum of a finite number of cyclic groups, it follows by 2.13 that A_{p_i} is likewise a direct sum of a finite number of cyclic groups. That is [dropping the subscript on p_i], $A_p = Z_1 + Z_2 + \cdots + Z_m$ for some m and where $O(Z_i) = (p^{\alpha_i})$ for $i = 1, 2, \cdots, m$. Let k be the maximum of the α_i. Then 4.1 implies that V contains an element of order (p^k). Suppose that V contains two independent elements of order (p). Then V contains a direct sum of two groups V_1 and V_2 where each V_i contains an element of order (p). By 4.3 $C_V(A) \geq C_{V_1}(A) + C_{V_2}(A)$, and this implies that $C_V(A_p)$ contains at least $2m$ independent elements of order (p). By 2.7 there are only m independent elements in A_p [and hence in $C_V(A_p)$ because of the condition of isomorphism] of order (p). Hence this contradiction implies that any two elements of order (p) in V are dependent, so that $(p^k, V) = (V, p'^k)$ is cyclic of order (p^k).

8.2. THEOREM. *Let A and V be abelian groups. Then $C_V(S) \sim S$ for every subgroup S of A if, and only if, the following two conditions hold:*

a). *A is a direct sum of a finite number of cyclic groups.*

b). *If A contains an element of order (r) which may be 0 or not, then $(V, r') = (r, V)$ is cyclic of order (r).*

Proof. The sufficiency is a consequence of 5. 3. Assume that $C_V(S) \sim S$ for every subgroup S of A, and let x be an arbitrary element of A. Then $B = \{x\}$ is cyclic of some order, say (r), where (r) may be 0 or not. By 5. 1 $C_V(A) \sim (V, r')$, and by hypothesis $C_V(B) \sim B$; hence $B \sim (V, r')$. Now $O(B) = (r)$, and it follows that $O(V, r')$ is also (r). This proves statement b).

As is indicated in the remark to 5. 3, condition b) implies that A cannot contain both elements of order 0 and also elements of order different from 0. The two cases are, then, considered separately.

If A contains no element of order 0, the theorem is a consequence of 8. 1. Suppose then that A contains only elements of order 0, and let x be any element different from 0 in A. Then $X = \{x\}$ is cyclic of order 0. Since by hypothesis $C_V(X) \sim X$, 5. 1 implies that $(V, 0') = V$ is also cyclic of order 0. Let B be a greatest independent subset of A. Then by hypothesis $C_V(\{B\}) \sim \{B\}$, and 6. 7 implies that B contains only a finite number of elements. Hence $r(A)$ is finite, and condition a) follows from 7. 2.

Theorems 8. 3 to 8. 6 inclusive do not hold when R is a field.

8. 3. Lemma. *Let A and V be abelian groups and assume that R is not a field. Then the conditions*

a). *$C_V(S) \sim S$ for every subgroup S of A,*

b). *$C_V(A/S) \sim A/S$ for every subgroup S of A,*

imply that A does not contain elements of order 0.

Proof. Assume that x is an element of order 0 in A. Since R is not a field, there is in R an element $n \neq 0$ which is not a unit. Put $\{nx\} = B$. Then A/B contains the element $B + x$ of order $(n) \neq 0$. Since by condition b) $C_V(A/B) \sim A/B$, 6. 1 implies that V contains an element v of order (n). Denote by T the cyclic group generated by x. Then 5. 1 shows there exists a character f of T in V such that $f(x) = v$. Hence f is of order $(n) \neq 0$. Then it is impossible that $C_V(T)$ be isomorphic to T, in contradiction to condition a). Hence no such element x exists.

8. 4. Theorem. *If A and V are abelian groups and R is not a field, then the following two statements are equivalent:*

1). a. *A is a direct sum of a finite number of cyclic groups of orders $\neq 0$,*

b. *If there is an element in* A *of order* (n), *then* $(V, n') = (n, V)$ *is cyclic of order* (n).

2). a. $S \sim C_V(S)$ *for every subgroup* S *of* A,
b. $A/S \sim C_V(A/S)$ *for every subgroup* S *of* A.

Proof. If 2) is satisfied, then 8.3 implies that A does not contain elements of order 0, and statement 1) is then a consequence of 8.1.

Assume 1) is satisfied. If A is a direct sum of a finite number of cyclic groups of orders different from 0, then 3.13 implies that every subgroup and every quotient group A/S of A also satisfies the same condition. Condition 1) b certainly holds for S and A/S since it holds for A. Statement 2) is then a direct consequence of 5.3.

8.5. LEMMA. *Let* A *and* V *be abelian groups and assume that* R *is not a field. Then the statement* $C_V(A/S) \sim A/S$ *for every subgroup* S *of* A *implies that* A *does not contain a cyclic quotient group of order* 0.

Proof. Assume that A/T for T a subgroup of A is cyclic of order 0. Then A/T is generated by some element $T + x$. Since R is not a field, there exists an element $r \neq 0$ in R which is not a unit. Suppose $r(T + x)$ generates A/T. Then there exists an element $\bar{r}$ in R such that $\bar{r}r(T + x) = T + x$. Since $T + x$ is of order 0, it follows that $\bar{r}r = 1$ which is impossible since r is not a unit. Hence $r \cdot (A/T) < A/T$. Therefore there is a subgroup L of A such that $T \leq L < A$ and $L/T = r \cdot (A/T)$. Clearly A/L is a cyclic group of order $(r) \neq 0$. Since $C_V(A/L) \sim A/L$, it follows by 6.1 that V contains an element v of order $(r) \neq 0$. Now since A/T is of order 0, $C_V(A/T) \sim (V, 0') = V$. But as just shown, V contains an element of order (r). Hence it is impossible that V be cyclic of order 0, and therefore $C_V(A/T) \nsim A/T$. This contradiction completes the proof.

8.6. THEOREM. *Let* A *and* V *be abelian groups and assume that* R *is not a field. Then* $C_V(A/S) \sim A/S$ *for every subgroup* S *of* A *if, and only if,*

a). *A is a direct sum of a finite number of cyclic groups of orders* $\neq 0$,

b). *If A contains an element of order* (r), *then* $(V, r') = (r, V)$ *is cyclic of order* (r).

Proof. The necessity follows from 3.13 and 5.3.

Assume that A contains elements of order 0. Then $A' = A/F(A)$ contains only elements [different from 0] of order 0. Let $b_1, b_2, \cdots$ be a greatest independent subset of A'. Define B as the closure of the group generated by the elements $b_2, b_3, \cdots$. Then any two elements of $D = A'/B$ are dependent, since every element of D depends on $B + b_1$. Furthermore, if $B' \doteq \{B + b_1\}$, then $A'' = D/B'$ has only elements of orders different from 0, and A'' can be expressed as a quotient group of A. Hence by the hypothesis of the theorem it follows that $C_V(A'') \sim A''$. Then by 8. 1 A'' is a direct sum of a finite number of cyclic groups of orders $\neq 0$. This implies that D is generated by a finite number of elements of order 0 [namely, the finite number of representatives of A'' and $B + b_1$]. Since any two elements of D are dependent, this fact by 3. 12 implies that D is cyclic of order 0. Now D can be represented as a cyclic quotient group A/S'' of A, and hence it follows by 8. 5 that $C_V(A/S'') \not\sim A/S''$ for some S'' contained in A. This contradicts the hypothesis of the theorem, and hence the assumption that A contains an element of order 0 is wrong; i. e., A does not contain elements of order 0. Since the condition $C_V(A/S) \sim A/S$ for every subgroup S of A implies that $C_V(A) \sim A$, the sufficiency now follows from 8. 1.

If the ring R is a field, the following theorems may be stated in order to extend the theory for this particular case. The proofs are not difficult and hence are omitted.

8. 7. THEOREM. *If A and V are abelian groups admitting the field R as operator-system, then $C_V(A) \sim A$ if, and only if,*

a). *The rank of A is finite;* b). *V is cyclic and different from* 0.

8. 8. COROLLARY. *Let A and V be abelian groups admitting the field R as operator-system. Then $C_V(S) \sim S$ for every subgroup S of A if, and only if,*

a). *The rank of A is finite;* b). *V is cyclic and different from* 0.

8. 9. COROLLARY. *Let A and V be abelian groups admitting the field R as operator-system. Then $C_V(A/S) \sim A/S$ for every subgroup S of A if, and only if,*

a). *The rank of A is finite;* b). *V is cyclic and different from* 0.

Note in particular that 8. 6 does not hold when R is a field; 8. 9 is its analogue in this special case. For ordinary abelian groups the ring R becomes the domain of ordinary integers. Hence the theory for ordinary abelian groups is included in Theorems 8. 1 to 8. 6 inclusive.

OKLAHOMA AGRICULTURAL AND MECHANICAL COLLEGE.

ON A STATISTICS OF THE RAMANUJAN SUMS.*

By Aurel Wintner.

The Ramanujan sums[1] $c_m(n)$ are generalizations of the Möbius function $\mu(m)$, to which they reduce when $n=1$. It is easy to see[2] that the prime number theorem can be formulated in purely statistical terms, by saying that the function $\mu(m)$ has an asymptotic distribution function. Hence, it is natural to ask what, if any, is the asymptotic distribution of the function $c_m(n)$ of m for varying values of n. The following considerations deal with this question and with its statistical implications. It turns out that an asymptotic distribution function (with respect to m) exists for every n. This will be shown by applying the fundamental Tauberian theorem of prime number theory infinitely often, that is, once for every n. In principle, it would be possible to start with the classical case, $n=1$, of prime number theory, and to reduce the case of an arbitrary $n>1$ to the case $n=1$ of the Möbius function, by using an identity of O. Hölder;[3] an identity which expresses the Ramanujan sums in terms of the functions of Euler and Möbius and the greatest common divisor of m and n, as follows:

$$c_m(n)\phi\left(\frac{m}{(m,n)}\right)=\phi(m)\mu\left(\frac{m}{(m,n)}\right).$$

However, for the actual computation of the statistical momenta or Fourier-Stieltjes transforms, the analytical procedure to be applied here appears to be more convenient. In this way it will be possible to calculate the probabilities represented by the asymptotic distributions involved.

The Ramanujan sums are the real integers[4]

$$c_m(n), \qquad (m=1,2,\cdots,\ n=1,2,\cdots),$$

defined by

$$(1)\qquad c_m(n)=\sum_{\omega}\omega^n,\text{ i. e., } c_m(n)=\sum_{\substack{0<r\leqq m\\(r,m)=1}}\exp\frac{2\pi inr}{m}=\sum_{\substack{0<r\leqq m\\(r,m)=1}}\cos\frac{2\pi nr}{m},$$

---

* Received November 26, 1940.

[1] S. Ramanujan, *Collected Papers*, Cambridge University Press (1927), pp. 179-199.

[2] Cf. p. 111 below.

[3] O. Hölder, "Zur Theorie der Kreisteilungsgleichung," *Lichtenstein Memorial Volume*, 1936, pp. 13-23.

[4] That the numbers (1) are integers, is implied by the irreducibility theorem of Kronecker.

where ω runs through the primitive m-th roots of unity. The number of such roots is $\phi(m)$, where ϕ is Euler's function. It is easy to see that

$$-n \leqq c_m(n) \leqq n. \tag{2}$$

If m is a prime power, $m = p^k$, then (1) reduces to

$$c_{p^k}(n) = \begin{cases} p^k - p^{k-1} & \text{if } l \geqq k, \\ -p^{k-1} & \text{if } l = k-1, \\ 0 & \text{if } 0 \leqq l \leqq k-2, \end{cases} \tag{3}$$

where $l = l_{pn}$ is the function of p and n defined as follows:

$$l = l_{pn}: \qquad p^l \mid n \text{ but } p^{l+1} \nmid n. \tag{4}$$

Let (1) be thought of as a function of m for every fixed n. For instance, the first function is

$$c_m(1) = \sum_\omega \omega; \text{ so that } c_m(1) = \mu(m), \tag{5}$$

where μ is the Möbius symbol. A fundamental property of all these functions of m is expressed by the fact that each of them is a multiplicative function, i. e., that

$$c_{m_1 m_2}(n) = c_{m_1}(n) c_{m_2}(n) \text{ whenever } (m_1, m_2) = 1. \tag{6}$$

This implies the relation

$$c_1(n) = 1 \tag{7}$$

(which is clear, of course, from (1) also).

Let $c_m(n)^j$, where $j = 1, 2, \cdots$, denote the j-th power, c^j, of $c = c_m(n)$. Since $c_m(n)^j$ is, in view of (6), a multiplicative function of m, it is clear from the order of magnitude of the values (3), that the Dirichlet series

$$f_n^{(j)}(s) = \sum_{m=1}^{\infty} \frac{c_m(n)^j}{m^s} \tag{8}$$

is absolutely convergent for $\sigma > 1$ and admits the Euler factorization

$$f_n^{(j)}(s) = \prod_p \left(1 + \sum_{k=1}^{\infty} \frac{c_{p^k}(n)^j}{p^{ks}} \right), \tag{9}$$

if $\sigma > 1$, where $s = \sigma + it$.

It is clear from (3) that

$$\sum_{k=1}^{\infty} \frac{c_{p^k}(n)^j}{p^{ks}} = \sum_{k=1}^{l} \frac{(p^k - p^{k-1})^j}{p^{ks}} + \frac{(-p^l)^j}{p^{(l+1)s}}, \text{ if } l \neq 0, \tag{10}$$

while

$$(11)\qquad \sum_{k=1}^{\infty} \frac{c_{p^k}(n)^j}{p^{ks}} = \frac{(-1)^j}{p^s}, \text{ if } l = 0,$$

where l is, according to (4), uniquely determined by p and n. Thus (9) can be reordered into

$$(12)\qquad f_n^{(j)}(s) = F_n^{(j)}(s) \prod_p \left(1 + \frac{(-1)^j}{p^s}\right), \qquad (\sigma > 1),$$

if $F_n^{(j)}(s)$ is defined by

$$(13)\quad F_n^{(j)}(s) = \prod_{p^g \mid n} \left(1 + \frac{(-1)^j}{p^s}\right)^{-1} \left(1 + \sum_{k=1}^{g} \frac{p^k - p^{k-1}}{p^{ks}} + \frac{(-p^g)^j}{p^{(g+1)s}}\right);$$

it being understood that the indices p, g of the last product run through those pairs of primes p and positive integers g for which p^g is a divisor of n, *where only the highest exponent is reckoned for every prime divisor p of n.* For instance, if n is a prime power, $n = p^k$, then (13) reduces to

$$(14)\quad F_{p^k}^{(j)}(s) = \left(1 + \frac{(-1)^j}{p^s}\right)^{-1} \left(1 + \sum_{k=1}^{k} \frac{(p^k - p^{k-1})^j}{p^{ks}} + \frac{(-1)^j p^{kj}}{p^{(k+1)s}}\right).$$

Hence, on writing k for g in (13), one sees that

$$(15)\qquad F_n^{(j)}(s) = \prod_{p^k \mid n} F_{p^k}^{(j)}(s)$$

for every n.

The functions (9) can be expressed in terms of the functions (14) and their products (15), as follows:

$$(16)\qquad f_n^{(2j)}(s) = F_n^{(2j)}(s)\zeta(s)/\zeta(2s),$$

while

$$(17)\qquad f_n^{(2j-1)}(s) = F_n^{(2j-1)}(s)/\zeta(s).$$

This is clear from (12), since, by the Euler factorization of the Riemann zeta function,

$$\prod_p \left(1 + \frac{1}{p^s}\right) = \frac{\zeta(s)}{\zeta(2s)}, \qquad \prod_p \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)}; \qquad (\sigma > 1).$$

It is clear that the functions (14) are regular in the half-plane $\sigma > 0$. Hence, (15) shows that the same is true of the functions $F_n^{(j)}(s)$. It follows, therefore, from (16), (17) and from the non-vanishing of $\zeta(s)$ on the line $\sigma = 1$, that the functions $f_n^{(j)}(s)/\zeta(s)$ attain continuous boundary values on the line $\sigma = 1$. Since (2) implies that, if n and j are fixed, $c_m(n)^j$ remains bounded as $m \to \infty$, it follows that Ikehara's Tauberian theorem is applicable to each of the Dirichlet series (8), which define the functions $f_n^{(j)}(s)$ for

$\sigma > 1$. Accordingly, there exists for every n and every j an average $M_m\{c_m(n)^j\}$, where

$$(18) \qquad M_m\{a_m\} = \lim_{m\to\infty} \frac{a_1 + \cdots + a_m}{m}.$$

According to Dirichlet's Abelian theorem, the value of $M_m\{c_m(n)^j\}$ must be the limit of $f_n^{(j)}(s)/\zeta(s)$ as $s \to 1$. Consequently, from (16),

$$(19) \qquad M_m\{c_m(n)^{2j}\} = F_n^{(2j)}(1)/\zeta(2),$$

while, from (17),

$$(20) \qquad M_m\{c_m(n)^{2j-1}\} = 0.$$

According to (15), the Abelian relation (19) can be written as

$$(21) \qquad M_m\{c_m(n)^{2j}\} = \frac{6}{\pi^2} \Pi_{p^k|n} F_{p^k}^{(2j)}(1),$$

while (14) shows that the factors of the product (21) are given by

$$(22) \qquad F_{p^k}^{(2j)}(1) = \left(1 + \frac{1}{p}\right)^{-1} (1 + \sum_{h=1}^{k} \frac{(p^h - p^{h-1})^{2j}}{p^h} + p^{(2j-1)k-1}).$$

Let the 0-th power of $c = c_m(n)$ be defined to be 1 also if $c = 0$:

$$(23) \qquad c_m(n)^j = 1 \text{ if } j = 0; \text{ so that } M_m\{c_m(n)^0\} = 1,$$

by (18).

It is clear from (2) and (18) that

$$|\, M_m\{c_m(n)^{2j}\} \,| = M_m\{c_m(n)^{2j}\} \leqq n^{2j}.$$

Hence, the radius of convergence of the power series

$$(24) \qquad G_n(u) = \sum_{j=0}^{\infty} \frac{M_m\{c_m(n)^{2j}\}}{(2j)!} (-u^2)^j$$

is infinite for every n $(= 1, 2, \cdots)$. According to (23) and (21), these entire functions can be written in the form

$$(25) \qquad G_n(u) = 1 + \frac{6}{\pi^2} \sum_{j=1}^{\infty} \frac{(-u^2)^j}{(2j)!} \Pi_{p^k|n} F_{p^k}^{(2j)}(1).$$

In particular, if n is a prime power, $n = p^k$, then, according to (22),

$$G_{p^k}(u) = 1 + \frac{6/\pi^2}{1 + p^{-1}} \sum_{j=1}^{\infty} (1 + \sum_{h=1}^{k} \frac{(p^h - p^{h-1})^{2j}}{p^h} + p^{(2j-1)k-1}) \frac{(-u^2)^j}{(2j)!}.$$

Since the last series can be written as the sum, $\sum_{j=1}^{\infty}$, of

$$\frac{\{-u^2\}^j}{(2j)!} + \sum_{h=1}^{k} \frac{1}{p^h} \frac{\{-u^2(p^h - p^{h-1})^2\}^j}{(2j)!} + \frac{1}{p^{k+1}} \frac{\{-u^2 p^{2k}\}^j}{(2j)!},$$

it is identical with

$$-1+\cos u+\sum_{h=1}^{k}\frac{-1+\cos\,(p^h-p^{h-1})u}{p^h}+\frac{-1+\cos p^k u}{p^{k+1}}.$$

Hence,

$$G_{p^k}(u)=1+\frac{6/\pi^2}{1+p^{-1}}\left(-\sum_{g=0}^{k+1}\frac{1}{p^g}+\cos u+\sum_{h=1}^{k}\frac{\cos\,(p^h-p^{h-1})u}{p^h}+\frac{\cos p^k u}{p^{k+1}}\right),$$

and so

$$(26)\qquad G_{p^k}(u)=\left(1-\frac{6}{\pi^2}\,\frac{1-p^{-k-2}}{1-p^{-2}}\right)+\frac{6}{\pi^2}\left(1+\frac{1}{p}\right)^{-1}\left(\sum_{h=0}^{k}\frac{\cos\phi(p^h)u}{p^h}+\frac{\cos p^k u}{p^{k+1}}\right)$$

where ϕ is Euler's function ($\phi(p^0)=1$).

It is clear from (26) that, if k is fixed,

$$(27)\qquad G_{p^k}(u)\to\left(1-\frac{6}{\pi^2}\right)+\frac{6}{\pi^2}\cos u\quad\text{as}\quad p\to\infty,$$

and that, if p is fixed,

$$(28)\qquad G_{p^k}(u)\to\left(1-\frac{6}{\pi^2(1-p^{-2})}\right)+\frac{6}{\pi^2}\left(1+\frac{1}{p}\right)^{-1}\sum_{h=0}^{\infty}\frac{\cos\phi(p^h)u}{p^h}\quad\text{as}\quad k\to\infty$$

provided that u is restricted to be real; and that (27) and (28) hold uniformly for $-\infty<u<\infty$.

According to (13),

$$(29)\qquad F_1^{(j)}(s)\equiv 1.$$

Hence, from (19),

$$(30)\qquad M_m\{c_m(1)^{2j}\}=6/\pi^2;\quad (j=1,2,\cdots)\qquad\text{[cf. (5)].}$$

Consequently, (21) can be written in the form

$$(31)\qquad M_m\{c_m(n)^{2j}\}=\prod_{p^k|n}{}^{*}\,F_{p^k}^{(2j)}(1),$$

where the asterisk means that $p^k=1$ is thought of as a divisor of n. On applying (31) to $n=p^k(\geqq 1)$, and substituting the result into (31), one sees that

$$(32)\qquad M_m\{c_m(n)^{2j}\}=\prod_{p^k|n}{}^{*}\,M_m\{c_m(p^k)^{2j}\}$$

for every $n\geqq 1$ and for every $j\geqq 0$.

Correspondingly, from (25) and (30),

$$(33)\qquad G_n(u)=1+\sum_{j=1}^{\infty}\frac{(-u^2)^j}{(2j)\,!}\prod_{p^k|n}{}^{*}\,F_{p^k}^{(2j)}(1)=\sum_{j=0}^{\infty}\frac{(-u^2)^j}{(2j)\,!}\prod_{p^k|n}{}^{*}\,F_{p^k}^{(2j)}(1),$$

by (23). Hence, from (31) and (32),

$$(34)\qquad G_n(j)=\sum_{j=0}^{\infty}\frac{(-u^2)^j}{(2j)\,!}\prod_{p^k|n}{}^{*}\,M_m\{c_m(p^k)^{2j}\}.$$

If $n=1$, then (34) reduces, in view of (30), to

$$G_1(u)=(1-6/\pi^2)+(6/\pi^2)\cos u \qquad \text{[cf. (27)]}. \tag{35}$$

A real sequence $a_1,\cdots,a_m,\cdots$ of numbers is said to have an asymptotic distribution function, ρ, if there exists a monotone function $\rho(\alpha)$, $-\infty<\alpha<\infty$, which satisfies the boundary conditions $\rho(-\infty)=0$, $\rho(\infty)=1$ and is such that

$$\frac{N_m(\alpha)}{m}\rightarrow\rho(\alpha); \qquad m\rightarrow\infty, \tag{36}$$

holds at every continuity point α of ρ, where $N_m(\alpha)$ denotes the number of those among the m values $a_1,\cdots,a_m$ which are less than α. It is known[5] that, if the sequence $\{a_m\}$ is bounded,

$$\limsup_{m\rightarrow\infty}|a_m|<\infty, \tag{37}$$

it has an asymptotic distribution function, ρ, if and only if the m-average $M_m\{a_m{}^j\}$ exists for every j; in which case

$$M_m\{a_m{}^j\}=[\rho]_j \quad \text{for} \quad j=0,1,2,\cdots; \quad ([\rho]_0=1), \tag{38}$$

where $[\rho]_j$ denotes the j-th momentum,

$$[\rho]_j=\int_{-\infty}^{\infty}\alpha^j d\rho(\alpha). \tag{39}$$

Obviously,

$$|[\rho]_j|\leqq(\limsup_{m\rightarrow\infty}|a_m|)^j. \tag{40}$$

As an illustration, let $a_m=\mu(m)$, where μ is the Möbius function. The prime number theorem is equivalent to the estimate

$$\mu(1)+\cdots+\mu(m)=o(m), \qquad m\rightarrow\infty,$$

which, in turn, is equivalent to the existence of the average $M_m\{\mu(m)\}$; it being an elementary fact that if this average exists, it must be 0 (Euler). Furthermore, it is an elementary fact that $M_m\{|\mu(m)|\}$ exists $(=6/\pi^2)$. Since $\mu(m)$ is either 0 or ± 1, it follows that the prime number theorem is equivalent to the statement that $M_m\{\mu(m)^j\}$ exists for every j, i.e., that $\mu(m)$ has an asymptotic distribution function, ρ. Clearly, $\rho=\rho(\alpha)$ is the step function which has the saltus $3/\pi^2$, $1-6/\pi^2$, $3/\pi^2$ at $\alpha=-1, 0, 1$, respectively.

---

[5] Cf., e.g., A. Wintner, "On an application of diophantine approximations to the repartition problems of dynamics," *Journal of the London Mathematical Society*, vol. 7 (1932), pp. 242-246.

Correspondingly, the above results on the Ramanujan sequences (1), where n is arbitrarily fixed, can be formulated in terms of asymptotic distribution functions, as follows:

Let $a_m = c_m(n)$, where n is fixed. According to (2), condition (37) is satisfied. Hence, the existence of the averages (19), (20) is equivalent to the statement that the function $a_m = c_m(n)$ of m has an asymptotic distribution function, $\rho_n = \rho_n(\alpha), -\infty < \alpha < \infty$. Furthermore, from (20), (30), (32) and (38),

$$(41) \qquad [\rho_n]_j = \prod_{p^k | n}^{*} [\rho_{p^k}]_j; \qquad (j = 0, 1, 2, \cdots),$$

where, according to (19), (22), (20) and (30)

$$[\rho_{p^k}]_{2j} = \frac{6}{\pi^2}\left(1+\frac{1}{p}\right)^{-1}\left(\sum_{h=0}^{k}\frac{\phi(p^h)^{2j}}{p^h} + p^{(2j-1)k-1}\right), \qquad [\rho_{p^k}]_{2j-1} = 0 \text{ if } k > 0$$

(42)

while $$[\rho_1]_{2j} = 6/\pi^2, \qquad [\rho_1]_{2j-1} = 0; \qquad (j = 1, 2, \cdots).$$

Let $L(u;\rho)$ denote the Fourier-Stieltjes transform of ρ; so that

$$(43) \qquad L(u;\rho) = \int_{-\infty}^{\infty} e^{iu\alpha} d\rho(\alpha); \qquad -\infty < u < \infty.$$

Thus, from (39) and (40), the power series expansion

$$(44) \qquad L(u;\rho) = \sum_{j=0}^{\infty} \frac{i^j[\rho]_j}{j!} u^j, \qquad (L(0;\rho) = [\rho]_0 = 1),$$

is valid for every u.

It is clear from (20), (24), (41) and (44) that

$$(45) \qquad G_n(u) = L(u;\rho_n).$$

As an application of the fundamental relation (45), one sees from (27) and (35) that

$$(46) \qquad \lim_{p\to\infty} L(u;\rho_{p^k}) = L(u;\rho_1), \text{ i. e., } \lim_{p\to\infty} \rho_{p^k}(\alpha) = \rho_1(\alpha),$$

where k is any fixed positive integer. A similar interpretation holds for (28).

According to (45) and (26),

$$(47)\qquad L(u;\rho_{p^k})=\left(1-\frac{6}{\pi^2}\,\frac{1-p^{-k-2}}{1-p^{-2}}\right)+\frac{6}{\pi^2}\left(1+\frac{1}{p}\right)^{-1}\text{ times}$$
$$\left(\sum_{h=0}^{k}\frac{\exp i\phi(h)u}{2p^h}+\frac{\exp ip^k u}{2p^{k+1}}+\sum_{h=0}^{k}\frac{\exp -i\phi(h)u}{2p^h}+\frac{\exp -ip^k u}{2p^{k+1}}\right).$$

Hence, it is seen from (43) that $\rho_{p^k}=\rho_{p^k}(\alpha)$, where $\rho_{p^k}(-\infty)=0$, is the step function which has at

$\alpha=0,\qquad\qquad \alpha=\pm\phi(p^h)$, where $h=1,\cdot\cdot\cdot,k,\qquad\qquad \alpha=\pm k$

the respective jumps

$$1-\frac{6}{\pi^2}\,\frac{1-p^{-k-2}}{1-p^{-2}},\qquad \frac{3}{\pi^2 p^h}\left(1+\frac{1}{p}\right)^{-1},\qquad \frac{3}{\pi^2 p^k}\left(1+\frac{1}{p}\right)^{-1}$$

(which add up to 1); while (35) shows again that $\rho_1=\rho_1(\alpha)$ has at $\alpha=0$ and $\alpha=\pm 1$ the jumps $1-6/\pi^2$ and $3/\pi^2$, respectively.

Accordingly, all that remains to be discussed is the statistical meaning of the operation which expresses an arbitrary distribution function ρ_n in terms of the distribution functions ρ_1 and ρ_{p^k}, where p^k is an arbitrary prime power.

To this end, consider first two distribution functions, say $\tau_1(\alpha)$ and $\tau_2(\alpha)$, which are constant for $-\infty<\alpha<0$ and are absolutely continuous; so that the distributions can be expressed in terms of their densities $\tau'_1(\alpha)$, $\tau'_2(\alpha)$, which vanish for $-\infty<\alpha<0$ (the primes denote differentiations with respect to α). Clearly, the Mellin transform,

$$(48)\qquad \int_0^\infty \tau'_1\left(\frac{\alpha}{\beta}\right)\frac{\tau'_2(\beta)}{\beta}\,d\beta,\qquad -\infty<\alpha<\infty,$$

defines a function which is the density of a distribution function; and this distribution function, which will be denoted by $\tau_1\dagger\tau_2\equiv\tau_1(\alpha)\dagger\tau_2(\alpha)$, satisfies the same conditions as either of its components τ_1, τ_2. It is also seen that

$$(49)\qquad \tau_1\dagger\tau_2=\tau_2\dagger\tau_1\quad\text{and}\quad(\tau_1\dagger\tau_2)\dagger\tau_3=\tau_1\dagger(\tau_2\dagger\tau_3),$$

where τ_3 is any distribution function satisfying the same conditions.

Clearly, the Mellin convolution $\tau_1\dagger\tau_2$ has with reference to the multiplication of independent functions the same statistical meaning as the usual convolution,

$$\tau_1*\tau_2\equiv\tau_1(\alpha)*\tau_2(\alpha)=\int_{-\infty}^{\infty}\tau_1(\alpha-\beta)\,d\tau_2(\beta),$$

has with reference to addition.

If τ_1 and τ_2 have finite momenta, the same is true of $\tau_1 \dagger \tau_2$; in which case it is easily verified from (39) and from the definition, (48), of the density of $\tau_1 \dagger \tau_2$, that

$$[\tau_1 \dagger \tau_2]_j = [\tau_1]_j [\tau_2]_j. \tag{50}$$

Obviously, the assumption that the τ are absolutely continuous is unessential in these considerations. In fact, it is sufficient that the τ should vanish for $-\infty < \alpha < 0$, and that a trivial precaution be taken for the possible jumps at $\alpha = 0$.

Consequently, the above explicit description of ρ_1 and of the ρ_{p^k}, when compared with (41) and (49)-(50), supplies an arbitrary ρ_n in explicit terms of statistical significance. In fact, ρ_1 and ρ_{p^k} are symmetric distribution functions and can, therefore, be reduced to the case of distribution functions which are constant for $-\infty < \alpha < 0$ and carry for $\alpha > 0$ a doubled probability.

THE JOHNS HOPKINS UNIVERSITY.

PROJECTIONS OF NON-ABELIAN GROUPS UPON ABELIAN GROUPS CONTAINING ELEMENTS OF INFINITE ORDER.*

By Ross A. Beaumont

I. Introduction. A function f of the subgroups of a group G upon the subgroups of a group G' is called a *projectivity* of G upon G' if:

(1) For every subgroup S of G, $f(S)$ is a subgroup of G';

(2) If S' is a subgroup of G', then there exists a subgroup S of G such that $f(S) = S'$;

(3) If S and T are subgroups of G, $S \leqq T$ is a necessary and sufficient condition that $f(S) \leqq f(T)$.

The correspondence f is a single-valued monotone-increasing function with a single-valued inverse and is a generalization of the ordinary concept of a projectivity between a projective space S_n and a projective space S'_n where for subspaces S_j and S_k of S_n, $S_j \leqq S_k$ is a necessary and sufficient condition that $S'_j \leqq S'_k$.

In the study of groups which have projections on abelian groups, it is natural to divide abelian groups into two classes: groups without elements of infinite order, and groups with elements of infinite order. In this paper the latter are studied. Baer has shown[1] that if an abelian group G contains at least two independent elements of infinite order, then every projectivity of G is induced by an isomorphism. Hence in a survey of groups which have projections on abelian groups containing elements of infinite order, only abelian groups G where $G/F(G)$ is of rank 1 need be considered. An important class of such groups are those for which $G/F(G)$ is an infinite cyclic group.

We first derive necessary conditions that a group have a projection on an abelian group G containing elements of infinite order; and, where $G/F(G)$ is an infinite cyclic group, we show that, under certain additional conditions, these necessary conditions are sufficient.

For future reference we collect the following definitions and notations:

A projectivity f of a group G upon a group G' is said to be *index-*

---

* Received March 10, 1941; Revised April 15, 1941; Presented to the American Mathematical Society, April 13, 1940.

[1] R. Baer, "The Significance of the System of Subgroups for the Structure of the Group," *American Journal of Mathematics*, vol. 61 (1939), pp. 1-44. Hereafter this paper will be referred to as (B).

The following formulae for multiplication of the elements in H are a consequence of 2.3. For x in $F(H,p)$ we have by complete induction from 2.3:

2.3.1. $$v^k x^m v^{-k} = x \text{ pow } m\,[1 + pc(p,v)]^k.$$

We then have by complete induction from 2.3.1:

2.3.2 $$\begin{aligned}(x^m v^k)^n &= x \text{ pow } [m(1 + \{1 + pc(p,v)\}^k + \{1 + pc(p,v)\}^{2k} \\ &\qquad + \cdots + \{1 + pc(p,v)\}^{(n-1)k})]\, v^{nk} \\ &= x \text{ pow } [m(\{1 + pc(p,v)\}^{nk} - 1)/(\{1 + pc(p,v)\}^k - 1)]\, v^{nk}.\end{aligned}$$

Now consider an element a of $F(H)$ where a is of composite order, i.e. $a = \prod_p a(p)$, $a(p)$ in $F(H,p)$ and almost every $a(p) = 1$.

$$\begin{aligned}([\prod_p a(p)]v^k)^n &= (\prod_{j=0}^{n-1} v^{kj}\,[\prod_p a(p)]\,v^{-kj})v^{nk} = (\prod_{j=0}^{n-1}\prod_p v^{kj}a(p)v^{-kj})v^{nk} \\ &= (\prod_{j=0}^{n-1}\prod_p a(p) \text{ pow } [1 + pc(p,v)]^{kj})v^{nk} \qquad \text{by 2.3.1} \\ &= (\prod_p \prod_{j=0}^{n-1} a(p) \text{ pow } [1 + pc(p,v)]^{kj})v^{nk} \\ &= (\prod_p a(p) \text{ pow } [1 + \{1 + pc(p,v)\}^k \\ &\qquad + \cdots + \{1 + pc(p,v)\}^{(n-1)k}])v^{nk}\end{aligned}$$

2.3.3. $$= (\prod_p a(p) \text{ pow } [(\{1 + pc(p,v)\}^{nk} - 1)/(\{1 + pc(p,v)\}^k - 1)])v^{nk}.$$

2.3.2 and 2.3.3 lead to the following useful formula:

2.3.4. $$([\prod_p a(p)]\, v^k)^n = (\prod_p [a(p)v^k]^n v^{-kn})v^{nk}.$$

We now derive other necessary conditions.

2.4. *If v is any element of infinite order in H, then* $c(2,v) \equiv 0 \bmod 2$ *if $F(G,2)$ contains an element of order* 4.

Proof. By Theorem 3.2 in (B) there exists an element z of infinite order in G such that $f(\{z\}) = \{v\}$; suppose u is in $F(G,2)$ where the order of u is at least 4. By Theorem 9.2 in (B) there exists one and only one element u' in H such that $f(\{u\}) = \{u'\}$ and $f(\{uz\}) = \{u'v\}$. By 2.2 $f(\{u^2\}) = f(\{u\})^2 = \{u'\}^2$ and $f(\{z^2\}) = f(\{z\})^2 = \{v^2\}$ and again by Theorem 9.2 in (B) there exists one and only one element u'' in H such that $f(\{u^2\}) = \{u''\}$ and $f(\{u^2z^2\}) = \{u''v^2\}$. Since $\{u''\} = \{u'^2\}$, $u'' = u'^{2\rho}$ where ρ is odd, and $f(\{u^2z^2\}) = \{u'^{2\rho}v^2\}$. By 2.2 $f(\{u^2z^2\}) = f(\{uz\}^2) = f(\{uz\})^2 = \{(u'v)^2\}$. Since u is in $F(G,2)$, $u \text{ pow } 2^k = 1$, $f(\{u \text{ pow } 2^k\}) = f(\{u\}) \text{ pow } 2^k = \{u' \text{ pow } 2^k\} = 1$ and u' is in $F(H,2)$. By 2.3.2 $(u'v)^2 =$

u' pow $[1 + \{1 + 2c(2, v)\}]\, v^2 = u'$ pow $2[1 + c(2, v)]\, v^2$. We now have $\{u'^{2\rho} v^2\} = \{u'$ pow $2\ [1 + c(2, v)] v^2\}$ which implies that $u'^{2\rho} = u'$ pow $2\,[1 + c(2, v)]$, $2\rho \equiv 2\,[1 + c(2, v)] \bmod 2^k$, $\rho \equiv 1 + c(2, v) \bmod 2^{k-1}$. Now $k > 1$, since the order of u is at least 4. Since ρ is odd, $1 + c(2, v)$ is odd and $c(2, v)$ is even. Hence $c(2, v) \equiv 0 \bmod 2$.

Later we shall need a particular form of the exponent $([1 + pc(p, v)]^{nk} - 1)/([1 + pc(p, v)]^{k} - 1)$ which appears in 2.3.2 for certain values of n and k. We prove the following lemma:

LEMMA 2.4.1. *If* $(i, p) = 1$ *and* $s \geqq t \geqq 0$, *then*

$([1 + pc(p, v)]^{ip^s} - 1)/([1 + pc(p, v)]^{p^{s-t}} - 1) = kp^t$ *where* $(k, p) = 1$.

Proof. We first prove a special case of the lemma:

2.4.1.1. $([1 + pc(p, v)]^{ip^s} - 1)/([1 + pc(p, v)] - 1) = mp^s$
where $(m, p) = 1$.

$$\begin{aligned}([1 + pc(p, v)]^{ip^s} - 1)/([(1 + pc(p, v)] - 1)\\ = (ip^s pc(p, v) + ip^s(ip^s - 1)p^2c(p, v)^2/2!\\ + \cdots + p \text{ pow } [ip^s]c(p, v) \text{ pow } [ip^s])/pc(p, v)\end{aligned}$$

2.4.1.2 $$\begin{aligned}= ip^s + ip^s(ip^s - 1)pc(p, v)/2!\\ + \cdots + p \text{ pow } [ip^s - 1]c(p, v) \text{ pow } [ip^s - 1].\end{aligned}$$

Consider the general term of the expression 2.4.1.2. It has the form:

$$ip^s(ip^s - 1)(ip^s - 2) \cdots (ip^s - r + 1)p^{r-1}c(p, v)^{r-1}/r!.$$

Now $r = mp^\tau$ where $m = 1$ or $(m, p) = 1$. We have:

$ip^s(ip^s - 1) \cdots (ip^s - mp^\tau + 1)/mp^\tau! = a$, an integer.

$(ip^s - 1)(ip^s - 2) \cdots (ip^s - mp^\tau + 1)/(mp^\tau - 1)! = b$, an integer. $a = ip^s b/mp^\tau$ and since m is prime to p, $ib/m = c$, an integer. Hence the general term has the form:

$cp^s p$ pow $[mp^\tau - 1]c(p, v)$ pow $[mp^\tau - 1]/p^\tau$ and we wish to show that cp pow $[mp^\tau - 1]c(p, v)$ pow $[mp^\tau - 1]/p^\tau$ is an integer divisible by p for $r = mp^\tau > 1$.

Case I $\tau = 0$

If $\tau = 0$ the above expression has the form: $cp^{m-1}c(p, v)^{m-1}$. $r = m > 1$ so that $cp^{m-1}c(p, v)^{m-1}$ is an integer divisible by p.

Case II $\tau > 0$

It is sufficient to prove that $\tau < mp^\tau - 1$ or $\tau + 1 < mp^\tau$. Now $\tau + 1$

$\leqq 2^\tau \leqq p^\tau \leqq mp^\tau$ and $2^\tau = p^\tau$ only when $p = 2$, since $\tau > 0$, so that we need only examine the case $p = 2$. $\tau + 1 = 2\tau$ only when $\tau = 1$, since $\tau > 0$. When $p = 2$, $\tau = 1$, the above expression has the form $c2^{2m-1}c(2,v)^{2m-1}/2$ which is an integer divisible by 2 for all $m > 1$. When $m = 1$, $c(2,v)$ is divisible by 2 if $F(G,2)$ contains an element of order 4 since $c(2,v) \equiv 0$ mod 2 by 2. 4. If all the elements of $F(G,2)$ are of order 2, it follows from 2. 3 that the elements of $F(H,2)$ permute with any element of infinite order in H, and hence 2. 3. 2 reduces to $(x^m v^k)^n = x^{mn} v^{nk}$ so that 2. 4. 1. 1 is trivially satisfied.

Hence every term after the first term of:

2. 4. 1. 3
$$i + i(ip^s - 1)pc(p,v)/2! \\ + \cdots + p \operatorname{pow}[ip^s - 1 - s]c(p,v) \operatorname{pow}[ip^s - 1]$$

is an integer divisible by p and since i is prime to p, the expression 2. 4. 1. 3 is an integer prime to p, which completes the proof of 2. 4. 1. 1.
Now

$$\begin{aligned} &([1 + pc(p,v)]^{ip^s} - 1)/([1 + pc(p,v)]^{p^{s-t}} - 1) \\ &\quad = ([1 + pc(p,v)]^{ip^s} - 1)/([1 + pc(p,v)] - 1) \\ &\qquad \div ([1 + pc(p,v)]^{p^{s-t}} - 1)/([1 + pc(p,v)] - 1) \\ &\quad = m_1 p^s / m_2 p^{s-t} = m_3 p^t \end{aligned}$$

where $(m_j, p) = 1$ by 2. 4. 1. 1, and this completes the proof of the lemma.

Remark. Note that for $p \neq 2$ every term after the first term of 2. 4. 1. 3 is divisible by $pc(p,v)$ so that 2. 4. 1. 2 has the form:

$$\begin{aligned} &p^s[i + pc(p,v)\{i(ip^s - 1)/2 \\ &\qquad + \cdots + p \operatorname{pow}[ip^s - s - 2]c(p,v) \operatorname{pow}[ip^s - 2]\}] \\ &\quad = p^s[\mathrm{i} + kpc(p,v)]. \end{aligned}$$

Then in the Lemma $m_1 = [i + Apc(p,v)]$, $m_2 = [1 + Bpc(p,v)]$, and $m_1/m_2 = k$, so that $k = i + (A - kB)pc(p,v) = i + xpc(p,v)$.

2. 5. *If $F(G,p)$ contains at least two independent elements of order p^n, then every v of infinite order in H permutes with every x in $F(H,p)$ such that x pow $p^n = 1$.*

Proof. If v is an element of infinite order in H, then $f(\{z\}) = \{v\}$ where $\{z\}$ is an infinite cyclic subgroup of G by Theorem 3. 2 (B). Suppose that $F(G)$ contains two independent elements, s, t, of order n and let $F(G)_n = (F(G);\ x^n = 1)$. Let $\bar{G} = F(G)_n \times \{z\}$ and $f(\bar{G}) = \bar{H}$. It follows from 2. 2 that $f(F(G)_n) = F(H)_n = (F(H);\ x^n = 1)$. Hence $f(\bar{G}) = \{f(F(G)_n),\ f(\{z\})\} = \{F(H)_n,\ v\} = \bar{H}$ and $F(H)_n$ is a characteristic subgroup of $\bar{H}$ and in particular is a normal subgroup. Now $\bar{G}^n = \{z^n\}$ and $f(\bar{G}^n) = f(\{z\}^n) = f(\{z\})^n = \{v^n\} = \bar{H}^n$ by 2.2. $\bar{H}^n$ is a characteristic subgroup of $\bar{H}$ and hence is normal in $\bar{H}$. Since $f(\bar{G}) = \bar{H}$ and $f(\bar{G}^n) = \bar{H}^n$, the lattice of subgroups between $\bar{G}$ and $\bar{G}^n$ is isomorphic to the lattice of subgroups between $\bar{H}$ and $\bar{H}^n$ and this lattice isomorphism induces a projectivity of $\bar{G}/\bar{G}^n$ upon $\bar{H}/\bar{H}^n$ which is index-preserving.

Now all the elements of $\bar{G}/\bar{G}^n$ have orders dividing n and there are at least three independent elements of order n, namely $\{z^n\}s$, $\{z^n\}t$, and $\{z^n\}z$. By Lemma 11.5 (B) $\bar{G}/\bar{G}^n$ and $\bar{H}/\bar{H}^n$ are isomorphic and hence $\bar{H}/\bar{H}^n$ is abelian. Then the commutator subgroup of $\bar{H}$ is contained in $\bar{H}^n$ so that $vxv^{-1}x^{-1}$ is in $\bar{H}^n$ for x in $\bar{H}$. If x is of finite order, $F(\{x,v\}) = \{x\}$ and x is normal in $\{x, v\}$. Hence $vxv^{-1}x^{-1} = x^{i-1}$, i. e. $vxv^{-1}x^{-1}$ is in $F(\bar{H}) = F(H)_n$. Thus for x in $\bar{H}$ of finite order $vxv^{-1}x^{-1}$ is in $\bar{H}^n \cap F(H)_n = 1$, and v permutes with every x in $F(H)_n$. This fact is equivalent to 2. 5.

2. 5 is equivalent to the following condition:

2. 5. 1. *If $F(G, p)$ contains two independent elements of order p^n, then* $c(p, v) \equiv 0 \bmod p^{n-1}$ *for any element v of infinite order in H.*

Proof. Suppose 2. 5 is satisfied. Then s, of order p^n in $F(H)_{p^n}$ permutes with v, i. e. $vsv^{-1} = s$. By 2. 3 $vxv^{-1} = x$ pow $[1 + pc(p, v)]$ for every x in $F(H, p)$ and since $F(H, p) \geqq F(H)_{p^n}$ which contains s, $vsv^{-1} = s$ pow $[1 + pc(p, v)]$. Hence $s = s$ pow $[1 + pc(p, v)]$, $1 \equiv 1 + pc(p, v) \bmod p^n$, $pc(p, v) \equiv 0 \bmod p^n$, $c(p, v) \equiv 0 \bmod p^{n-1}$.

If 2. 5. 1 is satisfied, $c(p, v) \equiv 0 \bmod p^{n-1}$ and hence $c(p, v) = p^{n-1}c'(p,v)$. For every x in $F(H, p)$, $vxv^{-1} = x$ pow $[1 + p^n c'(p, v)]$. For x in $F(H)_{p^n}$, $vxv^{-1} = x$ so that v permutes with x.

By 2. 1 $f(F(G)) = F(H)$ is the subgroup of all the elements of finite order in H and $F(H)$ is isomorphic to $F(G)$, so that in particular $F(H)$ is abelian. $F(G) = \prod_p F(G, p)$, $F(H) = \prod_p F(H, p)$ and the isomorphism of $F(G)$ upon $F(H)$ sends each $F(G, p)$ into the corresponding $F(H, p)$. Let $F'(G, p)$ be any subgroup of $F(G, p)$ and let $F'(H, p)$ be the corresponding subgroup of $F(H, p)$ under the isomorphism of $F(G)$ upon $F(H)$. We prove the following:

LEMMA 2. 6. *$F'(H, p)$ is either part of the central of H or else is the direct product of a group K' and a group U' where K' is part of the central of H, where U' is either cyclic or of type*[2] *p^∞, and where U' contains elements of an order exceeding the orders of all the elements in K'.*

---

[2] A group G is a group of type p^∞ if G is generated by an infinite sequence u_i of elements so that u_0 is of order p and $(u_{i+1})^p = u_i$ for every i.

Proof. If $F'(G,p)$ contains two independent elements of order p^n for all $n > 0$ where n is relevant for $F'(G,p)$, then it is a consequence of 2.5 and the fact that $F(H)$ is abelian that $F'(H,p)$ is a part of the central of H.

If $F'(G,p)$ contains two independent elements of order p^n for $0 < n \leqq k$ and elements of higher order such that any two elements of order p^n are dependent for $n > k$, then it is a standard result[3] for abelian groups that $F'(G,p) = K \times U$ where U is either a cyclic group of order p^{k+h} or a group of type p^∞ according as the orders of the elements of $F'(G,p)$ are bounded or not bounded. Moreover the orders of all the elements of K divide p^k since U contains elements of all higher orders which are independent of all the elements in K.

Since $F'(H,p)$ is isomorphic to $F'(G,p)$, $F'(H,p) = K' \times U'$ where $K' \leqq H_{p^k}$ and H_{p^k} is part of the central of H by 2.5. U' is cyclic or of type p^∞ and contains elements of an order exceeding p^k.

III. The Case of Cyclic G/F(G). Let A be an abelian group without elements of infinite order and Z be an infinite cyclic group $\{z\}$. Then $G = A \times Z$ is an abelian group containing elements of infinite order where $A = F(G)$ and $G/F(G)$ is an infinite cyclic group. If f is a projectivity of G upon a group H, then H satisfies, in addition to the conditions 2.1, 2.3, 2.4, and 2.5, the condition:

3.1. *$H/F(H)$ is an infinite cyclic group.*

Proof. $F(H)$ is a characteristic subgroup of H and in particular is a normal subgroup. Since $f(G) = f(A \times Z) = H$ and $f(A) = F(H)$ by 2.1, the set of subgroups between A and $A \times Z$ is isomorphic to the set of subgroups between $F(H)$ and H and this isomorphism induces a projectivity of $(A \times Z)/A$ upon $H/F(H)$. Now $(A \times Z)/A$ is an infinite cyclic group and it follows from Theorem 3.2 in (B) that $H/F(H)$ is an infinite cyclic group.

Hence if H is a projection of the abelian group $G = F(G) \times \{z\}$, H satisfies 2.1, 2.3, 2.4, 2.5, and 3.1. It will be the object of the remainder of this paper to show that if H is a group satisfying these necessary conditions, then there exists a projectivity, f, of G upon H.

If H is any group satisfying 2.1, 2.3, 2.4, 2.5, and 3.1, then $H = \{F(H), v\}$ where $F(H)$ is isomorphic to $F(G)$ by 2.1 and $\{v\}$ is an infinite cyclic group by 3.1. By 2.3 v induces an automorphism in $F(H)$ and in every $F(H,p)$ which sends each subgroup of $F(H)$ into itself.

---

[3] E. g. R. Baer, *American Journal of Mathematics*, vol. 49 (1937), pp. 99-101.

$F(G) = \prod_p F(G, p)$ since $F(G)$ is abelian and since $F(H)$ is isomorphic to $F(G)$, $F(H) = \prod_p F(H, p)$.

All prime numbers, p, are divided into two classes:

(A) those prime numbers, p, for which $F(G, p)$ contains two independent elements of order p^n for all $n > 0$ where n is relevant for $F(G, p)$;

(B) those prime numbers, p, for which $F(G, p)$ contains two independent elements of order p^n for $0 < n \leqq k$ and elements of higher order such that any two elements of order p^n are dependent for $n > k$.

THEOREM 3.2. *If $G = A \times Z$ where $A = F(G)$ is an abelian group without elements of infinite order and Z is an infinite cyclic group, and if the prime number p is not of class* (A), *then $F(G, p)$ does not contain elements of order p^2; then if f is a projectivity of G upon any group H, G and H are isomorphic groups.*

Proof. If $f(G) = H$, H satisfies 2.1, 2.3, 2.4, 2.5, and 3.1, and $H = \{F(H), v\}$ where $F(H)$ is isomorphic to A and $\{v\}$ is an infinite cyclic group. $A = F(G) = \prod_p F(G, p)$ and $F(H) = \prod_p F(H, p)$. If p is a prime of class (A), it follows from Lemma 2.6 that $F(H, p)$ is contained in the central of H. If p is not a prime of class (A), $F(H, p)$ does not contain elements of order p^2 since $F(G, p)$ and $F(H, p)$ are isomorphic. Then by 2.3, $vxv^{-1} = x \text{ pow } [1 + pv(p, v)] = x$ for x in $F(H, p)$ and $F(H, p)$ is in the central of H. Thus $F(H)$ is in the central of H and H is abelian. Hence $H = F(H) \times \{v\}$ where $F(H)$ is isomorphic to A and $\{v\}$ is isomorphic to $\{z\}$.

Remark. Hereafter we shall denote $c(p, v)$ by $c(p)$. Every element of infinite order in H has the form av^k where a is in $F(H)$. By 2.3 $(av^k)x(av^k)^{-1} = x \text{ pow } [1 + pc(p, av^k)]$ for x in $F(H, p)$. By 2.3.1 $(av^k)x(av^k)^{-1} = a(v^k x v^{-k})a^{-1} = ax \text{ pow}[1 + pc(p)]^k a^{-1} = x \text{ pow}[1 + pc(p)]^k$ since $F(H)$ is abelian. Hence $[1 + pc(p)]^k \equiv [1 + pc(p, av^k)] \bmod p^n$ for every n relevant for $F(H, p)$.

IV. A Correspondence between the Cyclic Subgroups of $G = F(G) \times \{z\}$ and a Group H Satisfying the Necessary Conditions.

Definition. A (1-1) correspondence f between the cyclic subgroups of a group G and the cyclic subgroups of a group H is called *index-preserving* if for every cyclic subgroup $\{x\}$ of G, $f(\{x\}^m) = f(\{x\})^m$.

We shall show that given an abelian group $G = F(G) \times \{z\}$ and any group H satisfying 2.1, 2.3, 2.4, 2.5, and 3.1 that such an index-preserving (1-1) correspondence f exists between the cyclic subgroups of G and those of H. We shall establish this correspondence by first considering the case where $F(G)$ is a primary abelian group $F(G,p)$. After obtaining this correspondence between the cyclic groups, we shall extend it to a projectivity of G upon H.

THEOREM 4.1. *If* $G = F(G,p) \times \{z\}$ *and* H *is any group satisfying* 2.1, 2.3, 2.4, 2.5, *and* 3.1, *then there exists an index-preserving* (1-1) *correspondence* f *between the cyclic subgroups of* G *and those of* H.

Proof. If p is of class (A), it follows from the proof of Theorem 3.2 that G and H are isomorphic groups and this isomorphism induces the correspondence f.

1) If p is of class (B) and the orders of the elements of $F(G,p)$ are bounded, then by 2.6 $G = K \times \{u\} \times \{z\}$ and $H = \{K' \times \{u'\}, v\}$ where K' is contained in the central of H and $\{u'\}$ is a cyclic group of order p^{k+h}. Any cyclic subgroup of G has the form $\{au^m z^n\}$ where a is in K, $m \geqq 0$, $n \geqq 0$. If $n \neq 0$, $\{au^m z^n\}$ can be put in the canonical form $\{au \operatorname{pow} [ijp^t]$ $z \operatorname{pow} [ip^s]\}$, where $(i,p) = 1$, $(j,p) = 1$ or $j = 0$, and where a; i, t, s are uniquely determined, j is uniquely determined mod the order of u^m. If $n = 0$, $\{au^m\}$ can be put in the form $\{bu \operatorname{pow} [jp^t]\}$ where b is in K, $j = 0$ or $j = 1$. b is not in this case uniquely determined. Conversely it is obvious that every a; i, j, t, s uniquely determine a cyclic subgroup of G.

Similarly every cyclic subgroup of H has either the form $\{a'u' \operatorname{pow} [ijp^t]$ $v \operatorname{pow} [ip^s]\}$ or the form $\{a'u' \operatorname{pow} [jp^t]\}$ where a'; i, j, t, s are determined as above.

We now establish a correspondence, f_1, which maps:

4.1.1. Case $i = 0$.

$$f_1(\{au \operatorname{pow} [jp^t]\}) = \{a'u' \operatorname{pow} [jp^t]\}$$

4.1.2. Case $i \neq 0$, $t \geqq s$.

$$\begin{aligned} &f_1(\{au \operatorname{pow} [ijp^t] z \operatorname{pow} [ip^s]\}) \\ &\quad = \{a'u' \operatorname{pow} [jp^{t-s}([1+pc(p)]^{ip^s} - 1)/([1+pc(p)] - 1)] v \operatorname{pow} [ip^s]\} \end{aligned}$$

which if $p \neq 2$ has the form

$$\{a'u' \operatorname{pow} jp^t[i + \psi(i,p,s)pc(p)]\, v \operatorname{pow} [ip^s]\}$$

by 2.4.1 and the remark following 2.4.1;

4.1.3. Case $i \neq 0$, $t < s$.

$$f_1(\{au \text{ pow } [ijp^t] z \text{ pow } [ip^s]\})$$
$$= \{a'u' \text{ pow } [j([1 + pc(p)]^{ip^s} - 1)/([1 + pc(p)]^{p^{s-t}} - 1)]\, v \text{ pow } [ip^s]\}$$

which if $p \neq 2$ has the form

$$\{a'u' \text{ pow } jp^t[i + \phi(p, i, s, t)pc(p)]\, v \text{ pow } [ip^s]\}$$

by 2.4.1 and the remark following 2.4.1; where $a \leftrightarrow a'$, $u \leftrightarrow u'$ under the isomorphism of $F(G, p)$ upon $F(H, p)$. In 4.1.2 the exponent of u' is given by 2.3.2 applied on $(u' \text{ pow } [jp^{t-s}]v) \text{ pow } [ip^s]$, i.e. with $m = jp^{t-s}$, $k = 1$, $n = ip^s$. In 4.1.3 the exponent of u' is given by 2.3.2 applied on $(u'^j v \text{ pow } [p^{s-t}]) \text{ pow } [ip^t]$, i.e. with $m = j$, $k = p^{s-t}$, $n = ip^t$.

If p is of class (B) and the orders of the elements of $F(G, p)$ are not bounded, then by 2.6, $G = K \times U \times \{z\}$ and $H = \{K' \times U', v\}$ where K' is contained in the central of H and U' is a group of type p^∞. It follows from the definition[2] of a group of type p^∞ that any cyclic subgroup of G has the form $\{au_k^m z^n\}$ where a is in K and u_k is the $(k+1)$-st generator in the infinite sequence of generators of U. u_k is of order p^{k+1}. If $n = 0$ the cyclic group has the form $\{bu_k \text{ pow } [jp^t]\}$ where if $k - t > 0$, $k - t$ is uniquely determined; $j = 0$ or $j = 1$; b is in K. If $n \neq 0$ the cyclic group has the form $\{au_k \text{ pow } [ijp^t] z \text{ pow } [ip^s]\}$ where $(i, p) = 1$; $(j, p) = 1$ or $j = 0$; if $k - t > 0$, a, i, $k - t$, s are uniquely determined; j is uniquely determined mod the order of u_k^m. Conversely every a; i, j, $k - t$, s uniquely determine a cyclic subgroup of G. Moreover; $au_k \text{ pow } [ijp^t] z \text{ pow } [ip^s] = au_{k-t}^{ij} z \text{ pow } [ip^s] = au_{k-t+s} \text{ pow } [ijp^s] z \text{ pow } [ip^s] = a(u^j_{k-t+s} z) \text{ pow } [ip^s]$. Note that if $y > x$, $u_{x-y} = 1$. Hence every infinite cyclic group of G has the form $\{a(u^j_{k-t+s} z) \text{ pow } [ip^s]\}$. Similarly every cyclic subgroup of H has either the form $\{a'u'_k \text{ pow } [jp^t]\}$ or $\{a'u'_{k-t+s} \text{ pow } [ijp^s] v \text{ pow } [ip^s]\}$ where a; i, j, $k - t$, s are determined as above.

We establish a correspondence, f_2, which maps:

4.1.4. $f_2(\{au_k \text{ pow } [jp^t]\}) = \{a'u'_k \text{ pow } [jp^t]\}$;

4.1.5. $f_2(\{a(u^j_{k-t+s} z) \text{ pow } [ip^s]\})$
$$= \{a'u'_{k-t+s} \text{ pow } [j([1 + pc(p)]^{ip^s} - 1)/([1 + pc(p)] - 1)]\, v \text{ pow } [ip^s]\}$$

which if $p \neq 2$ has the form

$$\{a'u'_k \text{ pow } jp^t[i + \psi(p, i, s)pc(p)]\, v \text{ pow } [ip^s]\}$$

by 2.4.1 and the remark following 2.4.1; where $a \leftrightarrow a'$, $u_{k-t+s} \leftrightarrow u'_{k-t+s}$ under the isomorphism between $F(G, p)$ and $F(H, p)$. Note that the mapping defined by 4. 1. 5 is essentially the same as in 4. 1. 2. The exponent of u'_{k-t+s} is obtained in the same way from 2. 3. 2.

Both mappings f_1 and f_2, defined above, are single-valued, since identical cyclic subgroups of G have the same canonical form which determines the mapping.

2) Suppose that two cyclic subgroups of G are mapped on the same cyclic subgroup of H by f_1 or f_2. If one of these cyclic subgroups of G is of finite order, the other must be also since both mappings send finite cyclic groups into finite cyclic groups and infinite cyclic groups into infinite cyclic groups. Hence both cyclic subgroups of G are finite or both are infinite. Since for finite cyclic groups both f_1 and f_2 are induced by the isomorphism of $F(G, p)$ upon $F(H, p)$, finite cyclic groups of G which are mapped on the same cyclic subgroup of H are identical.

If both cyclic subgroups of G are of infinite order they have the form

$$\{a_1 u \text{ pow } [i_1 j_1 p^{t_1}] z \text{ pow } [i_1 p^{s_1}]\} \quad \text{and} \quad \{a_2 u \text{ pow } [i_2 j_2 p^{t_2}] z \text{ pow } [i_2 p^{s_2}]\},$$

if the orders of the elements of $F(G, p)$ are bounded, and are mapped by f_1 into

$$\{a'_1 u' \text{ pow } [j_1 p^{t_1} m_1] v \text{ pow } [i_1 p^{s_1}]\} \quad \text{and} \quad \{a'_2 u' \text{ pow } [j_2 p^{t_2} m_2] v \text{ pow } [i_2 p^{s_2}]\}$$

where m_1 and m_2 are integers prime to p by 2. 4. 1. We are assuming

$$\{a'_1 u' \text{ pow } [j_1 p^{t_1} m_1] v \text{ pow } [i_1 p^{s_1}]\} = \{a'_2 u' \text{ pow } [j_2 p^{t_2} m_2] v \text{ pow } [i_2 p^{s_2}]\}.$$

Since v is of infinite order and $i_1 p^{s_1}$ and $i_2 p^{s_2}$ are positive,

$$a'_1 u' \text{ pow } [j_1 p^{t_1} m_1] v \text{ pow } [i_1 p^{s_1}] = a'_2 u' \text{ pow } [j_2 p^{t_2} m_2] v \text{ pow } [i_2 p^{s_2}].$$

Hence $i_1 p^{s_1} = i_2 p^{s_2}$, $i_1 = i_2$, $s_1 = s_2$; $a'_1 = a'_2$, $a_1 = a_2$; $j_1 m_1 p^{t_1} \equiv j_2 m_2 p^{t_2} \bmod p^{k+h}$, $t_1 = t_2$. Since $s_1 = s_2$, and $t_1 = t_2$, both cyclic groups in G are mapped by the same rule 4. 1. 2 or 4. 1. 3 and hence $m_1 = m_2$. Then $j_1 p^{t_1} \equiv j_2 p^{t_2} \bmod p^{k+h}$, $i_1 j_1 p^{t_1} \equiv i_2 j_2 p^{t_2} \bmod p^{k+h}$, $u \text{ pow } [i_1 j_1 p^{t_1}] = u \text{ pow } [i_2 j_2 p^{t_2}]$ and hence

$$a_1 u \text{ pow } [i_1 j_1 p^{t_1}] z \text{ pow } [i_1 p^{s_1}] = a_2 u \text{ pow } [i_2 j_2 p^{t_2}] z \text{ pow } [i_2 p^{s_2}]$$

so that the cyclic subgroups in G are identical.

If the orders of the elements of $F(G, p)$ are not bounded, the two infinite cyclic subgroups of G have the form

$$\{a_1 (u^{j_1}_{k_1-t_1+s_1} z) \text{ pow } [i_1 p^{s_1}]\} \quad \text{and} \quad \{a_2 (u^{j_2}_{k_2-t_2+s_2} z) \text{ pow } [i_2 p^{s_2}]\}$$

and are mapped by f_2 (4. 1. 5) into

$$\{a' u'_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] v \text{ pow } [i_1 p^{s_1}]\}$$

and

$$\{a'_2 u'_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}] v \text{ pow } [i_2 p^{s_2}]\}$$

which we are assuming identical. m_1 and m_2 are prime to p by 2. 4. 1. Again we have

$$a'_1 u'_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] v \text{ pow } [i_1 p^{s_1}] = a'_2 u'_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}] v \text{ pow } [i_2 p^{s_2}]$$

and hence: $i_1 p^{s_1} = i_2 p^{s_2}$, $i_1 = i_2$, $s_1 = s_2$; $m_1 = m_2$; $a'_1 = a'_2$, $a_1 = a_2$;

$$u'_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] = u'_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}].$$

If $k_1 = k_2$, then

$$u_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] = u_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}]$$

by the argument above, and hence the cyclic subgroups in G are identical. If $k_1 < k_2$ then $u'_{k_1} = u'_{k_2} \text{ pow } [p^{k_2-k_1}]$ and we have:

$$\begin{aligned} &u'_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] \\ &\qquad = u'^{\,j_1 m_1}_{k_1-t_1} = u'_{k_1} \text{ pow } [j_1 m_1 p^{t_1}] = u'_{k_2} \text{ pow } [j_1 m_1 p^{t_1+k_2-k_1}] \\ &\qquad = u'_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}] = u'^{\,j_2 m_2}_{k_2-t_2} = u'_{k_2} \text{ pow } [j_2 m_2 p^{t_2}]. \end{aligned}$$

Hence

$$j_1 m_1 p^{t_1+k_2-k_1} \equiv j_2 m_2 p^{t_2} \bmod p^{k_2+1}, \qquad j_1 p^{t_1+k_2-k_1} \equiv j_2 p^{t_2} \bmod p^{k_2+1},$$
$$u_{k_2} \text{ pow } [j_1 p^{t_1+k_2-k_1}] = u_{k_2} \text{ pow } [j_2 p^{t_2}], \qquad w^{j_1}_{k_1-t_1} = w^{j_2}_{k_2-t_2},$$
$$u_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] = u_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}].$$

If $k_2 < k_1$, $u'_{k_2} = u'_{k_1} \text{ pow } [p^{k_1-k_2}]$, and the above process shows

$$u_{k_1-t_1+s_1} \text{ pow } [j_1 m_1 p^{s_1}] = u_{k_2-t_2+s_2} \text{ pow } [j_2 m_2 p^{s_2}].$$

In either case the cyclic subgroups in G are identical.

We have shown that both mappings f_1 and f_2 send distinct cyclic subgroups of G into distinct cyclic subgroups of H.

3) Every cyclic subgroup of H has the form $\{a'u' \text{ pow } [jp^t]\}$ or $\{a'u' \text{ pow } [ijp^t] v \text{ pow } [ip^s]\}$ where the orders of the elements of $F(G, p)$ are bounded; the form $\{a'u'_k \text{ pow } [jp^t]\}$ or $\{a'u'_{k-t+s} \text{ pow } [ijp^s] v \text{ pow } [ip^s]\}$ where the orders of the elements of $F(G, p)$ are not bounded. For finite cyclic subgroups of H the inverse images under f_1 and f_2 are $\{au \text{ pow } [jp^t]\}$ and $\{au_k \text{ pow } [jp^t]\}$ respectively. Since

$$f_1(\{au \text{ pow } [ij'p^t] z \text{ pow } [ip^s]\}) = \{a'u' \text{ pow } [j'mp^t] v \text{ pow } [ip^s]\}$$

and

$$f_2(\{au_{k-t+s} \text{ pow } [ij'p^s] z \text{ pow } [ip^s]\}) = \{a'u'_{k-t+s} \text{ pow } [j'mp^s] v \text{ pow } [ip^s]\}$$

where $(m, p) = 1$ by 2. 4. 1, it is only necessary to choose j' so that $j'm \equiv ji$ mod p^r where p^r is the order of u' pow $[p^t]$ or u'_{k-t+s} pow $[p^s]$. This is possible since m, j, and i are all prime to p. Hence every infinite cyclic subgroup of H has an inverse image under f_1 and f_2.

In 1), 2), and 3) we have shown that each mapping f_1 and f_2, is a (1-1) correspondence.

4) Since for finite cyclic subgroups, both mappings f_1, and f_2 are induced by an isomorphism, f_1 and f_2 are index-preserving as far as finite cyclic subgroups are concerned.

If the orders of the elements of $F(G,p)$ are bounded, any cyclic subgroup of G has the form $\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\}$ and

$$\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\}^n = \{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\} \text{ pow } [mp^r]$$

where $(m,p)=1$. Suppose $t \geqq s$. Then by 4.1.2:

$$f_1(\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\}) = \{a'(u' \text{ pow } [jp^{t-s}]v) \text{ pow } [ip^s]\}.$$

$$\begin{aligned} &f_1(\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\} \text{ pow } [mp^r]) \\ &\quad = f_1(\{a \text{ pow } [mp^r]u \text{ pow } [ijmp^{t+r}]z \text{ pow } [imp^{s+r}]\}) \\ &\quad = \{a' \text{ pow } [mp^r](u' \text{ pow } [jp^{t-s}]v) \text{ pow } [imp^{s+r}]\} \\ &\quad = \{a'(u' \text{ pow } [jp^{t-s}]v) \text{ pow } [ip^s]\} \text{ pow } mp^r \\ &\quad = f_1(\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\}) \text{ pow } [mp^r] \end{aligned}$$

since a' is in K' which is contained in the central of H.

The above argument holds for $t < s$ since by 4.1.3

$$f_1(\{au \text{ pow } [ijp^t]z \text{ pow } [ip^s]\}) = \{a'(u'^j v \text{ pow } [p^{s-t}]) \text{ pow } [ip^t]\}.$$

Hence f_1 is index-preserving.

If the orders of the elements of $F(G,p)$ are not bounded, since

$$f_2(\{au_{k-t+s} \text{ pow } [ijp^s]z \text{ pow } [ip^s]\}) = \{a'(u'^j_{k-t+s}\, v) \text{ pow } [ip^s]\}$$

by 4.1.5, the same argument shows f_2 to be index-preserving.

We have shown that both mappings f_1 and f_2 are index-preserving (1-1) correspondences between the cyclic subgroups of G and those of H and this completes the proof of Theorem 4.1.

Again consider the abelian group $G = A \times Z$ where A contains no elements of infinite order and Z is an infinite cyclic group $\{z\}$.

$$G = F(G) \times \{z\} = \prod_p F(G,p) \times \{z\}$$

and

$$\prod_p F(G,p) = \prod_{p(A)} F(G,p(A)) \prod_{p(B_1)} F(G,p(B_1)) \prod_{p(B_2)} F(G,p(B_2))$$

where the primes $p(A)$ are primes of class A, $p(B_1)$ are primes of class B and the orders of the elements of $F(G,p(B_1))$ are bounded, $p(B_2)$ are primes of class B and the orders of the elements of $F(G,p(B_2))$ are not bounded.

THEOREM 4. 2. *If $F(G) \times \{z\}$ is an abelian group and H is any group satisfying* 2. 1, 2. 3, 2. 4, 2. 5, and 3. 1, *then there exists an index-preserving* (1-1) *correspondence f between the cyclic subgroups of G and those of H.*

Proof. $H = \{F(G), v\}$ where $F(H)$ is isomorphic to $F(G)$ and $\{v\}$ is an infinite cyclic group. It follows from Lemma 2. 6 that every element of G has the form:

$$\begin{aligned}&\prod_{p(A)} a(p) \prod_{p(B_1)} a(p)u(p)^{m(p)} \prod_{p(B_2)} a(p)u(p)_{k(p)}^{n(p)}\, z^r \\ &\qquad = \prod_{p} a(p) \prod_{p(B_1)} u(p)^{m(p)} \prod_{p(B_2)} u(p)_{k(p)}^{n(p)}\, z^r\end{aligned}$$

where $\prod\limits_{p} a(p)$ is mapped by the isomorphism of $F(G)$ upon $F(H)$ into the element $\prod\limits_{p} a'(p)$ which is in the central of H by Lemma 2. 6; $u(p)$ is of order a power of p; and $u(p)_{k(p)}$ is a canonical generating element of a group of type p^{∞}. We can write the general element in either the form:

4. 2. 1.
$$\begin{aligned}&\prod_{p} a(p) \times \prod_{p(B_1)[t(p) \geqq s(p)]} u(p) \text{ pow } [i(p)j(p)p^{t(p)}] \\ &\qquad \times \prod_{p(B_1)[t(p) < s(p)]} u(p) \text{ pow } [i(p)j(p)p^{t(p)}] \\ &\qquad \times \prod_{p(B_2)} u(p)_{k(p)-t(p)+s(p)} \text{ pow } [i(p)j(p)p^{s(p)}] \times z^r\end{aligned}$$

where $0 < r = i(p)p^{s(p)}$, and each term of the product is reduced to the above form by the methods of Theorem 4. 1; or the form:

4. 2. 2.
$$\prod_{p} a(p) \prod_{p(B_1)} u(p) \text{ pow } [j(p)p^{t(p)}] \prod_{p(B_2)} u(p)_{k(p)} \text{ pow } [j(p)p^{t(p)}]$$

where $r = 0$.

Every cyclic subgroup of G is generated by an element of the form 4. 2. 1 or 4. 2. 2 and we establish a correspondence f which maps the cyclic subgroups generated by 4. 2. 1 and 4. 2. 2 upon the cyclic subgroups of H generated respectively by:

4. 2. 3.
$$\begin{aligned}&\prod_{p} a'(p) \\ &\times \prod_{p(B_1)[t(p) \geqq s(p)]} u'(p) \text{ pow } [j(p)p^{t(p)-s(p)}([1+pc(p)]^{i(p)p^{s(p)}} \\ &\qquad\qquad -1)/([1+pc(p)]-1)] \\ &\times \prod_{p(B_1)[t(p) < s(p)]} u'(p) \text{ pow } [j(p)([1+pc(p)]^{i(p)p^{s(p)}} \\ &\qquad\qquad -1)/([1+pc(p)]^{p^{s(p)-t(p)}}-1)] \\ &\times \prod_{p(B_2)} u'(p)_{k(p)-t(p)+s(p)} \text{ pow } [j(p)([1+pc(p)]^{i(p)p^{s(p)}} \\ &\qquad\qquad -1)/([1+pc(p)]-1)] \times v^r,\end{aligned}$$

4. 2. 4.
$$\prod_{p} a'(p) \prod_{p(B_1)} u'(p) \text{ pow } [j(p)p^{t(p)}] \prod_{p(B_2)} u'(p)_{k(p)} \text{ pow } [j(p)p^{t(p)}],$$

where the primed elements are the corresponds in H under the isomorphism of $F(G)$ upon $F(H)$ and the exponents of the $u'(p)$ are given by 4. 1. 2, 4. 1. 3, and 4. 1. 5.

f defines a (1-1) correspondence between the cyclic subgroups of G and those of H since the terms of the products 4. 2. 3 and 4.2. 4 which are in different primary groups $F(G, p)$ are independent, and each of these terms is mapped by a rule which gives a (1-1) correspondence by Theorem 4. 1.

We may now write 4. 2. 3 in another form:

4. 2. 5.
$$\begin{aligned}&\prod_{p} a'(p)\\ &\quad\times \prod_{p(B_1)[t(p)\geqq s(p)]} \{(u'(p) \text{ pow } [j(p)p^{t(p)-s(p)}]v) \text{ pow } [i(p)p^{s(p)}]\\ &\qquad\qquad v \text{ pow } [-i(p)p^{s(p)}]\}\\ &\quad\times \prod_{p(B_1)[t(p)<s(p)]} \{(u'(p)^{j(p)}v \text{ pow } [p^{s(p)-t(p)}]) \text{ pow } [i(p)p^{t(p)}]\\ &\qquad\qquad v \text{ pow } [-i(p)p^{s(p)}]\}\\ &\times \prod_{p(B_2)} \{(u'(p)^{j(p)}_{k(p)-t(p)+s(p)}\, v) \text{ pow } [i(p)p^{s(p)}]v \text{ pow } [-i(p)p^{s(p)}]\} \times v^{r}\end{aligned}$$

Now the n-th power of the cyclic subgroup of G generated by 4. 2. 1 is mapped by f on the cyclic subgroup of H generated by:

4. 2. 6.
$$\begin{aligned}&\prod_{p} a'(p)^{n}\\ &\quad\times \prod_{p(B_1)[t(p)\geqq s(p)]} \{(u'(p) \text{ pow } [j(p)p^{t(p)-s(p)}]v) \text{ pow } [ni(p)p^{s(p)}]\\ &\qquad\qquad v \text{ pow } [-ni(p)p^{s(p)}]\}\\ &\quad\times \prod_{p(B_1)[t(p)<s(p)]} \{(u'(p)^{j(p)}v \text{ pow } [p^{s(p)-t(p)}]) \text{ pow } [ni(p)p^{t(p)}]\\ &\qquad\qquad v \text{ pow } [-ni(p)p^{s(p)}]\}\\ &\times \prod_{p(B_2)} \{(u'(p)^{j(p)}_{k(p)-t(p)+s(p)}\, v) \text{ pow } [ni(p)p^{s(p)}]v \text{ pow } [-ni(p)p^{s(p)}]\} \times v^{nr}\end{aligned}$$

which is just the n-th power of 4. 2. 5 by 2. 3. 4.

Hence as far as infinite cyclic subgroups are concerned f is index-preserving. Since for finite cyclic groups f is induced by an isomorphism, f is also index-preserving on finite cyclic groups.

V. Extension of the Correspondence between the Cyclic Subgroups to a Projectivity of *G* upon *H*. Consider the following conditions on the correspondence, f, established in Theorem 4. 2, between the cyclic subgroups of G and those of H.

5. 1. If $S = \{a, b\} = \{a\} \times \{b\}$ is a subgroup of G where a is any element of finite order in G and b is any element of infinite order in G, then $f(\{ab\}) \leqq \{f(\{a\}), f(\{b\})\}$ in H.

5. 2. If $S' = \{\{c'\}, \{d'\}\}$ is a subgroup of H where c' is any element of finite order in H and d' is any element of infinite order in H, then

$$f^{-1}(\{c'd'\}) \leqq \{f^{-1}(\{c'\}), f^{-1}(\{d'\})\} = f^{-1}(\{c'\}) \times f^{-1}(\{d'\})$$

in G.

We shall show that if f satisfies 5.1 and 5.2, then there exists a projectivity of the abelian group $G = F(G) \times \{z\}$ which induces f. We shall further show under what conditions f satisfies 5.1 and 5.2.

If S is any subgroup of $G = F(G) \times \{z\}$, $F(S) = F(G) \cap S$ is a characteristic subgroup of S, and it is a standard result that $S/F(S)$ is isomorphic to a subgroup of $G/F(G)$. Then $S/F(S)$ is either the identity $F(S)$ or an infinite cyclic group $\{F(S)w\}$. In the latter case $S = F(S) \times \{w\}$ where w in S is of infinite order and hence $w = az^k$, $k > 0$, a is in $F(G)$. Thus every subgroup S of G has the form $F(S) \times \{w\}$ where either $w = 1$ or $w = az^k$, $k > 0$, a is in $F(G)$.

THEOREM 5.3. *If the* (1-1) *index-preserving correspondence, f, between the cyclic subgroups of G and those of H satisfies conditions* 5.1 *and* 5.2, *then the correspondence $g(S) = g(F(S) \times \{w\}) = \{F(S)', f(\{w\})\}$, where $F(S)'$ is the subgroup of H which corresponds to $F(S)$ under the isomorphism of $F(G)$ upon $F(H)$, is an index-preserving projectivity of the group G upon the group H.*

Proof. For every subgroup S of G, $g(S)$ is a subgroup of H and any subgroup S' of H has the form $\{F(S'), \{w'\}\}$ so that there exists a subgroup $S = F(S) \times f^{-1}(\{w'\})$ where $F(S)' = F(S')$ such that $g(S) = S'$.

5.3.1. If $S = F(S) \times \{w\}$, then $f(\{aw^h\}) \leqq \{F(S)', f(\{w\})\}$ where a is in $F(S)$.

$$\{F(S)', f(\{w\})\} \geqq f(\{w\})^h = f(\{w\}^h) = f(\{w^h\})$$
$$\{F(S)', f(\{w\})\} \geqq f(\{a\}) = \{a'\}.$$

Hence $\{F(S)', f(\{w\})\} \geqq \{f(\{w^h\}), f(\{a\})\}$. Now $S \geqq \{w^h\} \times \{a\}$ and by 5.1 $f(\{aw^h\}) \leqq \{f(\{w^h\}), f(\{a\})\}$.

Now consider $S_1 = F(S_1) \times \{w_1\} \leqq S_2 = F(S_2) \times \{w_2\}$. Since $F(S_1) \leqq F(S_2)$, $F(S_1)' \leqq F(S_2)'$. Now $w_1 = aw_2^h$ where a is in $F(S_2)$ and it follows from 5.3.1 that

$$f(\{w_1\}) = f(\{aw_2^h\}) \leqq \{F(S_2)', f(\{w_2\})\}.$$

Hence

$$\{F(S_1)', f(\{w_1\})\} \leqq \{F(S_2)', f(\{w_2\})\}$$

or $S_1 \leqq S_2$ implies $g(S_1) \leqq g(S_2)$. We have now shown that g is single-valued.

5.3.2. If $g(S) = \{F(S)', f(\{w\})\}$, then $f^{-1}(\{a'w'^h\}) \leqq F(S) \times \{w\}$, where a' is in $F(S)'$ and $\{w'\} = f(\{w\})$.

$$S = F(S) \times \{w\} \geqq \{w\}^h = f^{-1}(\{w'\})^h = f^{-1}(\{w'^h\})$$
$$S = F(S) \times \{w\} \geqq \{a\} = f^{-1}(\{a'\})$$

Hence $S \geqq f^{-1}(\{w'^h\}) \times f^{-1}(\{a'\})$ and by 5.2 $f^{-1}(\{a'w'^h\}) \leqq f^{-1}(\{w'^h\}) \times f^{-1}(\{a'\}) \leqq S$.

Consider $g(S_1) = \{F(S_1)', \ f(\{w_1\})\} \leqq g(S_2) = \{F(S_2)', \ f(\{w_2\})\}$. Since $F(S_1)' \leqq F(S_2)'$, $F(S_1) \leqq F(S_2)$. Let $f(\{w_1\}) = \{w'_1\}$ and $f(\{w_2\}) = \{w'_2\}$. $w'_1 = a'w'^h_2$ where a' is in $F(S_2)'$. It follows from 5.3.2 that $\{w_1\} = f^{-1}(\{w'_1\}) = f^{-1}(\{a'w'^h_2\}) \leqq S_2$. Hence $S_1 = F(S_1) \times \{w_1\} \leqq S_2 = F(S_2) \times \{w_2\}$.

We have shown that g satisfies the definition of a projectivity of G upon H. Since g induces f on the cyclic subgroups, g is index-preserving.

THEOREM 5.4. *If $G = F(G, p) \times \{z\}$ is an abelian group where $p \neq 2$ and H is any group satisfying* 2.1, 2.3, 2.4, 2.5, *and* 3.1, *then the correspondence, f, between the cyclic subgroups of G and those of H satisfies conditions* 5.1 *and* 5.2.

Proof. If p is a prime of class A, it follows from Theorem 3.2 that f is induced by an isomorphism of G upon H and hence 5.1 and 5.2 are satisfied.

5.4.1. *Proof of* 5.1.

Case I. p is a prime of class B_1.

By 2.6, $F(G, p) = K \times \{u\}$ where $\{u\}$ is a cyclic group of order p^{k+h} and the orders of all the elements of K divide p^k. $F(H, p) = K' \times \{u'\}$ where K' is contained in the central of H. Any element of finite order in G has the form $a = du \text{ pow } [gp^w]$ where d is in K, $(g, p) = 1$ or $g = 0$. Any infinite cyclic subgroup of G is generated by an element of the form $b = cu \text{ pow } [ijp^t]z \text{ pow } [ip^s]$ where c is in K, $(i, p) = 1$, $(j, p) = 1$ or $j = 0$.

Since $(i, p) = 1$, $(i, p^r) = 1$ where p^r is the order of a and there exist integers m, n such that $im + p^r n = 1$. $im \equiv 1 \bmod p^r$ where $(m, p) = 1$. Hence we have $a = a^{im} = d^{im}u \text{ pow } [imgp^w] = du \text{ pow } [imgp^w]$ and $ab = dcu \text{ pow } [(mgp^w + jp^t)i]z \text{ pow } [ip^s]$.

If $t < w$, $ab = dcu \text{ pow}[(gmp^{w-t} + j)ip^t]z \text{ pow } [ip^s]$ where $(gmp^{w-t} + j)$ is prime to p and we have by either 4.1.2 or 4.1.3:

$$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)A\}]v \text{ pow } [ip^s]\}$$
$$f(\{ab\}) = \{d'c'u' \text{ pow } [(gmp^{w-t} + j)p^t\{i + pc(p)A\}]v \text{ pow } [ip^s]\}$$

where $A = \phi(p, i, s, t)$ if $t < s$ and $A = \psi(p, i, s)$ if $t \geqq s$.

If $t > w$, $ab = dcu \text{ pow } [(gm + jp^{t-w})ip^w]z \text{ pow } [ip^s]$ where $(gm + jp^{t-w})$ is prime to p and there are three possibilities:

1) $s > t > w$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\phi(p, i, s, t)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 3.
$f(\{ab\}) = \{d'c'u' \text{ pow } [(gm + jp^{t-w})p^w\{i + pc(p)\phi(p, i, s, w)\}]v \text{ pow } [ip^s]\}$
by 4. 1. 3.

2) $t \geqq s > w$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 2
$f(\{ab\}) = \{d'c'u' \text{ pow } [(gm + jp^{t-w})p^w\{i + pc(p)\phi(p, i, s, w)\}]v \text{ pow } [ip^s]\}$
by 4. 1. 3.

3) $t > w \geqq s$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 2
$f(\{ab\}) = \{d'c'u' \text{ pow } [(gm + jp^{t-w})p^w\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\}$
by 4. 1. 2.

If $t = w$, $ab = dcu \text{ pow } [(gm + j)ip^t]z \text{ pow } [ip^s]$ and $gm + j = np^v$ where $(n, p) = 1$ or $n = 1$. Hence $ab = dcu \text{ pow } [inp^{t+v}]z \text{ pow } [ip^s]$ and there are three possibilities:

1) $t \geqq s$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\psi(p, i, s)\}] v \text{ pow } [ip^s]\}$ by 4. 1. 2
$f(\{ab\}) = \{d'c'u' \text{ pow } [np^{v+t}\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 2

2) $t + v \geqq s > t$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\phi(i, p, s, t)\}] v \text{ pow } [ip^s]\}$ by 4. 1. 3
$f(\{ab\}) = \{d'c'u' \text{ pow } [np^{v+t}\{i + pc(p)\psi(i, p, s)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 2

3) $s > t + v \geqq t$

$f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)\phi(i, p, s, t)\}] v \text{ pow } [ip^s]\}$ by 4. 1. 3
$f(\{ab\}) = \{d'c'u' \text{ pow } [np^{v+t}\{i + pc(p)\phi(i, p, s, t + v)\}]v \text{ pow } [ip^s]\}$ by 4. 1. 3.

Every possibility is enumerated above and each is included in one of the two general classifications:

(I. a) $f(\{b\}) = \{c'u' \text{ pow } [jp^t\{i + pc(p)A\}]v \text{ pow } [ip^s]\} = \{b_f\}$,
$f(\{ab\}) = \{d'c'u' \text{ pow } [(gmp^w + jp^t)\{i + pc(p)A\}]v \text{ pow } [ip^s]\}$.

By 4. 1. 1 $f(\{a\}) = \{d'u' \text{ pow } [gp^w]\} = \{a_f\}$ and we have:

$\{f(\{a\}), f(\{b\})\} = \{a_f, b_f\} \geqq \{a_f \text{ pow } [m\{i + pc(p)A\}]b_f\}$
$= \{(d'u' \text{ pow}[gp^w]) \text{ pow}[m\{i + pc(p)A\}]\, c'u' \text{ pow}[jp^t\{i + pc(p)A\}]$
$v \text{ pow}[ip^s]\}$
$= \{d'c'u' \text{ pow}[(mgp^w + jp^t)\{i + pc(p)A\}]v \text{ pow } [ip^s]\} = f(\{ab\})$

since $im \equiv 1 \bmod p^r$, $pc(p) \equiv 0 \bmod p^k$, and d' in K' is of order p^{k^*} where $k^* \leqq r$, $k^* \leqq k$.

$$\text{(I. b)} \qquad \begin{aligned} f(\{b\}) &= \{c'u' \text{ pow } [jp^t\{i + pc(p)A\}]v \text{ pow } [ip^s]\}, \\ f(\{ab\}) &= \{c'd'u' \text{ pow } [(mgp^w + jp^t)\{i + pc(p)B\}]v \text{ pow } [ip^s]\}, \end{aligned}$$

where $A \neq B$.

It should be noted that for every possibility in this classification $t \geqq w$. We shall prove that there exists an exponent y such that $\{a_f{}^y b_f\} = f(\{ab\})$ and hence $f(\{ab\}) \leqq \{f(\{a\}), f(\{b\})\}$.

Let $y = 1 + xp^{k^*}$ where p^{k^*} is the order of d' and hence $k^* \leqq k$. Then we have

$$a_f{}^y = (d'u' \text{ pow } [gp^w]) \text{ pow } [1 + xp^{k^*}] = d'u' \text{ pow } [gp^w + xgp^{w+k^*}]$$

and we wish to determine x so that:

$$\begin{aligned} &(d'u' \text{ pow } [gp^w + xgp^{w+k^*}])(c'u' \text{ pow } [jp^t\{i + pc(p)A\}]v \text{ pow } [ip^s]) \\ &\qquad = d'c'u' \text{ pow } [(mgp^w + jp^t)\{i + pc(p)B\}]v \text{ pow } [ip^s]. \end{aligned}$$

This is equivalent to finding an integral solution for the congruence:

$$\begin{aligned} &gp^w + xgp^{w+k^*} + jp^t\{i + pc(p)A\} \equiv (mgp^w + jp^t)\{i + pc(p)B\} \bmod p^{k+h}, \\ &xgp^{w+k^*} \equiv migp^w - gp^w + ijp^t - ijp^t \\ &\qquad\qquad\qquad\qquad + p^w pc(p)\{mgB + jp^{t-w}B - jp^{t-w}A\} \bmod p^{k+h}, \\ &xgp^{k^*} \equiv mig - g + pc(p)D \bmod p^{k+h-w}. \end{aligned}$$

Since $mi \equiv 1 \bmod p^r$ and $r \geqq k + h - w$ we have $xgp^{k^*} \equiv pc(p)D \bmod p^{k+h-w}$. Since $pc(p) \equiv 0 \bmod p^k$, $pc(p)D = p^{k^*}S$ and the congruence reduces to: $xg \equiv S \bmod p^{k+h-w-k^*}$. If $k^* \geqq k + h - w$, any value for x satisfies the congruence. Otherwise $x = nS$ where $ng \equiv 1 \bmod p^{k+h-w-k^*}$ is an integral solution of the congruence.

Case II. p is a prime of class B_2.

By 2. 6, $F(G, p) = K \times U$ where U is a group of type p^∞ and the orders of all the elements of K divide p^k. $F(H, p) = K' \times U'$ where K' is contained in the central of H. Every element of finite order in G has the form $a = du_{r_1} \text{ pow } [gp^w]$ where d is in K and u_{r_1} is one of the canonical generating elements of U. Every infinite cyclic subgroup of G is generated by an element of the form $b = cu_{r_2} \text{ pow } [ijp^t]z \text{ pow } [ip^s]$ where c is in K and u_{r_2} is a canonical generating element of U.

If $r_1 \leqq r_2$, $u_{r_1} = u_{r_2} \text{ pow } [p^{r_2-r_1}]$ and $a = du_{r_2} \text{ pow } [gp^{w+r_2-r_1}]$

Hence: $f(\{a\}) = \{d'u'_{r_2} \text{ pow } [gp^{w+r_2-r_1}]\} = \{a_f\}$ by 4. 1. 4

$f(\{b\}) = \{c'u'_{r_2} \text{ pow } [jp^t\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\} = \{b_f\}$ by 4. 1. 5.

m is determined so that $im \equiv 1 \bmod p^r$ where p^r is the order of a, and

$$a = a^{im} = (du_{r_2} \text{ pow } [gp^{w+r_2-r_1}])^{im} = du_{r_2} \text{ pow } [gimp^{w+r_2-r_1}]$$

and hence:

$$ab = dcu_{r_2} \text{ pow } [(mgp^{w+r_2-r_1} + jp^t)i]z \text{ pow } [ip^s]$$

and

$$f(\{ab\}) = \{d'c'u'_{r_2} \text{ pow } [(mgp^{w+r_2-r_1} + jp^t)\{i + pc(p)\psi(p, i, s)\}]v \text{ pow } [ip^s]\}$$

by 4. 1. 5.

By (I. a)

$$f(\{ab\}) = \{a_f \text{ pow } [m\{i + pc(p)\psi(p, i, s)\}]b_f\} \leqq \{f(\{a\}), f(\{b\})\}.$$

If $r_1 > r_2$, $u_{r_2} = u_{r_1}$ pow $[p^{r_1-r_2}]$ and the above argument again applies.

5. 4. 2. *Proof of* 5. 2.

Case I. p is a prime of class B_1.

$F(H, p) = K' \times U' = \{u'\}$ is a cyclic group of order p^{k+h}, K' is contained in the central of H, and the orders of all the elements of K' divide p^k. Any element of finite order in H has the form $a' = d'u'$ pow $[gp^w]$ where d' is in K', $(g, p) = 1$ or $g = 0$. Any infinite cyclic subgroup of H is generated by an element of the form $b' = c'u'$ pow $[ijp^t]v$ pow $[ip^s]$ where $(i, p) = 1$, $(j, p) = 1$ or $j = 0$, c' is in K'.

m is determined so that $im \equiv 1 \bmod p^r$ where p^r is the order of a'. Then $a' = d'u'$ pow $[gimp^w]$ and

$$a'b' = c'd'u' \text{ pow } [(gmp^w + jp^t)i]v \text{ pow } [ip^s].$$

(I. a) $t < w$

$$a'b' = c'd'u' \text{ pow } [(gmp^{w-t} + j)ip^t]\, v \text{ pow } [ip^s]$$

where $(gmp^{w-t} + j)$ is prime to p. It follows from 4. 1. 2 or 4. 1. 3 that:

$$f^{-1}(\{b'\}) = \{cu \text{ pow } [rijp^t]z \text{ pow } [ip^s]\} = \{b\}$$

where $r[i + pc(p)A] \equiv i \bmod p^{k+h}$.

Then $r(gmp^{w-t} + j)[i + pc(p)A] \equiv i(gmp^{w-t} + j) \bmod p^{k+h}$ and by 4. 1. 2 or 4. 1. 3:

$$\begin{aligned} f^{-1}(\{a'b'\}) &= \{cdu \text{ pow } [r(gmp^{w-t} + j)ip^t]z \text{ pow } [ip^s]\} \\ &= \{du \text{ pow } [rgimp^w]b\} = \{du \text{ pow } [rgp^w]b\} \end{aligned}$$

since $im \equiv 1 \bmod p^r$ and $r \geqq k + h - w$.

$$f^{-1}(\{a'\}) = \{du \operatorname{pow}[gp^w]\} = \{a\}$$

by 4. 1. 1.

Now $ri \equiv i - rpc(p)A \bmod p^{k+h}$, $i(r-1) \equiv -rpc(p)A \bmod p^{k+h}$ and since $(i,p) = 1$ and $pc(p) \equiv 0 \bmod p^k$ it follows that $r \equiv 1 \bmod p^k$ or $r = 1 + Bp^k$. Then $du \operatorname{pow}[rgp^w]b = (du \operatorname{pow}[gp^w])^r b$ since the order of d divides p^k. Hence

$$f^{-1}(\{a'b'\}) = \{a^r b\} \leqq \{f^{-1}(\{a'\}), f^{-1}(\{b'\})\}.$$

(I. b) $t \geqq w$

$$\begin{aligned} a'b' &= c'd'u' \operatorname{pow}[(gm + jp^{t-w})ip^w]v \operatorname{pow}[ip^s] \\ &= c'd'u' \operatorname{pow}[ij'p^{w+v}]v \operatorname{pow}[ip^s] \end{aligned}$$

where $(j', p) = 1$ and $j'p^v = gm + jp^{t-w}$. It follows from 4. 1. 2 or 4. 1. 3 that:

$$\begin{aligned} f^{-1}(\{a'b'\}) &= \{cdu \operatorname{pow}[si + j'p^{w+v}]z \operatorname{pow}[ip^s]\} \\ &= \{cdu \operatorname{pow}[si(gmp^w + jp^t)]z \operatorname{pow}[ip^s]\} \end{aligned}$$

where $s[i + pc(p)B] \equiv i \bmod p^{k+h}$.

We shall prove that there exists an exponent $y = 1 + xp^{k^*}$ where p^{k^*} is the order of d so that $\{a^y b\} = f^{-1}(\{a'b'\})$ and hence $f^{-1}(\{a'b'\}) \leqq \{f^{-1}(\{a'\}), f^{-1}(\{b'\})\}$. Thus we must determine x so that:

$$\begin{aligned} (du \operatorname{pow}[gp^w]) \operatorname{pow}[1 + xp^{k^*}](cu \operatorname{pow}[rijp^t]z \operatorname{pow}[ip^s]) &= cdu \operatorname{pow}[gp^w + gxp^{w+k^*} + rijp^t]z \operatorname{pow}[ip^s] \\ &= cdu \operatorname{pow}[si(gmp^w + jp^t)]z \operatorname{pow}[ip^s]. \end{aligned}$$

This is equivalent to finding an integral solution for the congruence:

$$\begin{aligned} &gp^w + gxp^{w+k^*} + rijp^t \equiv sigmp^w + sijp^t \bmod p^{k+h}, \\ &gxp^{k^*} \equiv simg - g + sijp^{t-w} - rijp^{t-w} \bmod p^{k+h-w}. \\ &gxp^{k^*} \equiv g(s-1) + ijp^{t-w}(s-r) \bmod p^{k+h-w} \end{aligned}$$

since $im \equiv 1 \bmod p^r$ and $r \geqq k + h - w$. Now $s \equiv 1 \bmod p^k$ and $r \equiv s \bmod p^k$ so that the congruence reduces to $gx \equiv gp^{k-k^*}M + ijp^{t-w+k-k^*}N \bmod p^{k+h-w-k^*}$, $gx \equiv S \bmod p^{k+h-w-k^*}$.

If $k^* \geqq k + h - w$, any value for x satisfies the congruence. Otherwise $x = nS$ where $ng \equiv 1 \bmod p^{k+h-w-k^*}$ is an integral solution.

Case II. p is a prime of class B_2.

The proof reduces to (I. a) above in the same way that it did in 5. 4. 1 and this completes the proof of Theorem 5. 4.

The above proof depends entirely on the remark following Lemma 2. 4. 1 which holds when $p \neq 2$. However, if $p = 2$ we may prove the following:

THEOREM 5.5. *If* $G = F(G,2) \times \{z\}$ *is an abelian group and* H *is any group satisfying* 2.1, 2.3, 2.4, 2.5, *and* 3.1, *then the correspondence,* f, *between the cyclic subgroups of* G *and those of* H *satisfies* 5.1 *and* 5.2 *if either*:

(a) 2 *is a prime of class* (A); *or*
(b) *all the elements of* $F(G,2)$ *have orders dividing* 4.

Proof. If (a) is satisfied, it follows from Theorem 4.2 that f is induced by an isomorphism of G upon H, and hence 5.1 and 5.2 are satisfied.

Suppose (b) is satisfied. Then if $F(G,2)$ contains an element of order 4, H is abelian by 2.4 and 2.3 and f is induced by an isomorphism of G upon H. If $F(G,2)$ does not contain an element of order 4, H is abelian by 2.3, and again f is induced by an isomorphism. Hence 5.1 and 5.2 are satisfied.

COROLLARY 5.6. *If* $G = F(G) \times \{z\}$ *is an abelian group which satisfies either*:

(a) 2 *is a prime of class* (A);
(b) *all the elements of* $F(G,2)$ *have orders dividing* 4; *or*
(c) 2 *is an irrelevant prime number for* G;

and H *is any group satisfying* 2.1, 2.3, 2.4, 2.5, *and* 3.1, *then the correspondence,* f, *between the cyclic subgroups of* G *and those of* H *satisfies* 5.1 *and* 5.2.

Proof. Since $G = \prod_p F(G,p) \times \{z\}$, every element of finite order in G has the form $\prod_p a(p)$ where $a(p)$ is in $F(G,p)$ and every element of infinite order has the form $[\prod_p b(p)]z^k$ where $b(p)$ is in $F(G,p)$. The orders of the factors of $\prod_p a(p)$ and of $\prod_p b(p)$ are relatively prime so that it follows from Theorems 5.4 and 5.5 and from the definition of the correspondence f given in 4.2.3 and 4.2.4 that f satisfies 5.1 and 5.2.

Corollary 5.6 and Theorem 5.3 give us our final result:

THEOREM 5.7. *If* $G = F(G) \times \{z\}$ *is an abelian group which satisfies* (a), (b), *or* (c) *of Corollary* 5.6, *then there exists a projectivity of* G *upon the group* H *if, and only if,* H *satisfies* 2.1, 2.3, 2.4, 2.5, *and* 3.1.

UNIVERSITY OF WASHINGTON,
SEATTLE, WASHINGTON,

HOMOGENEITY OF PROJECTIVE PLANES.*[1]

REINHOLD BAER.

A satisfactory analytic theory may be developed for every projective plane in which Desargues' Theorem is valid, since it is then possible to describe the points of the plane by means of coördinates taken from a (not necessarily commutative) field; and this field will be commutative if, and only if, Pappus' Theorem is valid in the plane. It is well known[2] that these properties are intimately connected with properties of the transformations the plane may undergo. But whereas Desargues' property is equivalent to the existence of certain transformations, the theorem of Pappus is usually linked with non-existence, i. e., with the property that there exists at most one transformation meeting certain requirements.

It is the object of this note to show that Pappus' and Desargues' properties are equivalent to conditions assuring the existence of transformations only. Since any point-point-transformation one might wish for exists in a projective plane over a (not necessarily commutative) field, we shall have to consider dualities which map points on lines and lines on points.

To every projective plane there exists a dual one and thus the theory as a whole is self-dual. But a projective plane over a field is self-dual if, and only if, this field admits of an anti-automorphism; and the identity is an anti-automorphism if, and only if, the field is commutative. Thus the conditions with which we are concerned will have to assure the existence of very special dualities, namely dualities possessing an axis and a center; as a matter of fact our main result states that it is sufficient to assure the existence of two one-parameter families of such dualities.

1. Preliminaries. A projective plane consists of two classes of elements: points $p, q, \cdots$, and lines $R, S, \cdots$ which are connected by one relation: $p < R$ signifying that the point p is on the line R and that the line R passes through the point p. The negation of this relation is denoted by $p \nless R$. Points, lines and the relation "$<$" are subject to the following rules.

---

* Received February 28, 1941.

[1] Presented to the American Mathematical Society, April, 1941.

[2] Cf. O. Veblen-J. W. Young: *Projective Geometry*. E. Artin: *Reports of a Mathematical Colloquium*, Second Series, 2 (1940), 15-20 and in particular W. Schwan: *Mathematische Zeitschrift* 3 (1919), 11-28.

(1) There exist at least two different lines and two different points.

(2) If p and q are two different points, then there exists one and only one line passing through p and q; this line is denoted by $p + q$.

(3) If P and Q are two different lines, then there exists one and only one point which is on both these lines; this point is denoted by PQ.

(4) Every point is on at least three different lines and every line contains at least three different points.

Points lying on a line and lines passing through one point shall be termed collinear points or lines respectively.

2. Dualities and transformations. A function f defined for all the points and lines of a projective plane is said to be a *transformation of the plane,* if it satisfies the following conditions.

(i) f is a $1-1$-correspondence of the class of all the points upon itself.

(ii) f is a $1-1$-correspondence of the class of all the lines upon itself.

(iii) $p^f + q^f = (p + q)^f, \qquad P^fQ^f = (PQ)^f.$

A function f defined for the points and lines of a projective plane is said to be a *duality of the plane,* if it satisfies the following conditions.

(i*) f is a $1-1$-correspondence which maps the class of points upon the class of all the lines.

(ii*) f is a $1-1$-correspondence which maps the class of lines upon the class of all the points.

(iii*) $p^fq^f = (p + q)^f, \qquad P^f + Q^f = (PQ)^f.$

The product of two dualities of the projective plane under discussion is clearly a transformation of the plane; and the transformations of the plane form a group. There exists always at least one transformation of the plane; but there are planes which do not admit of dualities.

Our discussion will not be concerned with all the possible transformations and dualities of the plane. If p is a point and L a line of the plane, then the transformation f of the plane is termed a $p-L$*-transformation,* if

(iv) $R = R^f$ *for* $p < R$ *and* $q = q^f$ *for* $q < L$.

If $p \nless L$, then the duality f of the plane is termed a $p-L$*-duality,* if

(iv*) $RL = R^f$ *for* $p < R$ *and* $q + p = q^f$ *for* $q < L$.

Again it is clear that the $p-L$-transformations form a group, that the inverse of a $p-L$-duality as well as the product of a $p-L$-duality and a $p-L$-transformation is a $p-L$-duality, and that the product of two $p-L$-dualities is a $p-L$-transformation.

Suppose now that f is a $p-L$-transformation, that q is a point which is different from p and which does not lie on L and that $q = q^f$. If R is a line, then R is certainly left invariant by f, if $R = L$ or if $p < R$. If $q < R$, then $R = q + RL$ so that such a line R is left invariant by f too, since both q and RL are fixed-points of f. If x is a point not on $p + q$, then $x = (p + x)(q + x)$. Hence such a point x is a fixed-point too, since both lines $p + x$ and $q + x$ are left invariant by f. Now it is clear that every line is left invariant by f, since a line is either equal to $p + q$ or else contains at least two different points not on $p + q$; and since all the lines are left invariant by f, all the points are fixed-points, i. e., $f = 1$. Likewise we might have proved that the $p-L$-transformation f is the identity, if it leaves invariant a line different from L not passing through p.

Noting that the quotients of two $p-L$-transformations as well as the quotients of two $p-L$-dualities are $p-L$-transformations, the following facts are direct consequences of what we have shown just now.

THEOREM 2. 1. *If f and g are $p-L$-transformations ($p-L$-dualities) such that $q^f = q^g$ for some point q different from p and not on L, then $f = g$.*

If f and g are $p-L$-transformations ($p-L$-dualities) such that $Q^f = Q^g$ for some line Q different from L and not containing p, then $f = g$.

The following criteria will prove helpful.

LEMMA 2. 2. *If R, S, L are three different lines, and if f is a transformation of the plane such that $R = R^f$, $S = S^f$ and $p = p^f$ for every $p < L$, then f is an $(RS)-L$-transformation.*

Proof. If the point RS is not on L, X any line passing through RS, then $X = RS + XL$; and $(RS)^f = R^f S^f = RS$ and $(XL)^f = XL$ imply $X = X^f$. Hence we assume now that $RS < L$, and that p is a point neither on R nor on S such that $p \neq p^f$. Then $p + p^f = P$ is a well determined line different from L so that $P = p + p^f = p^f + PL = PL + p$ and consequently $P = P^f$. All we have to prove now is that $PL = RS$. Let us assume that $PL \neq RS$. Then PR and PS are two different points not on L and they are fixed-points, since the lines P, R, S are left invariant by f. Hence it follows as before that every line through PR and every line through PS is left invariant by f. Thus it follows that f is both a $PR-L$- and a $PS-L$-transformation; and since PR and PS are different points not on L, it follows from Theorem 2. 1 that $f = 1$, a contradiction which completes the proof.

COROLLARY 2. 3. *If a transformation f of the plane leaves invariant all the points of a line L, then f is a p — L-transformation (for some point p).*

Proof. If $f \neq 1$, then there exists a point q, not on L, such that $q \neq q^f$. Since $(q + q^f)L$ is a fixed-point under f, the line $q + q^f$ is left invariant by f. Since there exist points that are neither on L nor on $q + q^f$, it follows that there exist at least two lines different from L that are left invariant by f; and now our statement is a consequence of Lemma 2. 2.

COROLLARY 2. 4. *If K and L are two (distinct or coincident) lines, then the p — L-transformations for $p < K$ form a group.*

Proof. Suppose that f is a p — L-transformation, that g is a q — L-transformation, and that both p and q are on K. It follows from Corollary 2. 3 that fg is an r — L-transformation for some point r. If $K \neq L$, $fg \neq 1$, then the line K is invariant under fg and r is on K, since otherwise $fg = 1$. If $K = L$, $fg \neq 1$, but r not on L, then $r^f = r^{g^{-1}}$ so that $p = q$ and $f = g^{-1}$ by Theorem 2. 1, contradicting $fg \neq 1$. Thus it follows that fg is an r — L-transformation for $r < K$; and this proves our statement.

3. Transitivity and homogeneity. The projective plane is p — L-*transitive,* if there exists, to every pair of points r, s such that r and s are different from p and not on L and such that p, r, s are collinear, a p — L-transformation mapping r upon s.

THEOREM 3. 1. *The projective plane is p — L-transitive, if there exists a point r different from p and not on L with the following property.*

If s is a point on $p + r$, not on L and different from p, then there exists a p — L-transformation mapping r upon s.

Proof. Suppose that u and v are two points (which are not on L and which are both different from p) such that p, u, v are collinear. Then either both u and v are on $p + r$ or neither of them is on $p + r$. In the first case there exists a p — L-transformation f mapping r upon u and a p — L-transformation g mapping r upon v so that the p — L-transformation $f^{-1}g$ maps u upon v. In the second case we consider the line $u + r$ different from both $p + r$ and L. Then $r = (u + r)(p + r)$, $L(u + r)$ is a point different from p; and $(p + r)(v + L(u + r)) = w$ is a point on $p + r$, not on L and different from p. There exists, therefore, a p — L-transformation f mapping r upon w and f maps u upon v, since

$$u = (u + p)(u + r) = (u + p)(r + L(u + r))$$

and

$$v = (v + p)(v + L(u + r)) = (u + p)(w + L(u + r));$$

and this completes the proof.

If K and L are two lines, then the projective plane is said to be *K — L-transitive,* if it is p — L-transitive for every point p on K.

LEMMA 3.2. *If K and L are two (equal or different) lines, and if the projective plane is p — L-transitive for at least two different points p on K, then the plane is K — L-transitive.*

Proof. Suppose that p and q are two different points on K such that the plane is both p — L- and q — L-transitive. If u and v are two different points which are neither on L nor on K, then $p + u$ and $q + v$ are two different lines such that $w = (p + u)(q + v)$ is a point not on K. We distinguish two cases.

A. w not on L.

Since $w \neq p$, and since p, u, w are collinear, there exists a p — L-transformation f, mapping u upon w; and since $w \neq q$, and q, v, w are collinear, there exists a q — L-transformation g, mapping w upon v. The product fg is a transformation of the plane which leaves every point on L invariant and which maps u upon v. Hence, it follows from Corollary 2.4 that fg is a $K(u + v)$ — L-transformation, mapping u upon v.

B. w on L.

Since our lemma clearly is true, if lines contain exactly three or four points, we may assume now that lines contain more than four points. Thus there exists on the line $u + v$ a point t which is different from u, v and which is neither on K nor on L. Since $t \neq v$ is on $u + v$, it follows that the two different lines $p + u$ and $q + t$ meet in a point not on L; and hence it follows from what has been proved in A, that there exists a $K(u + t)$ — L-transformation r mapping u upon t. Likewise it follows that there exists a $K(t + v)$ — L-transformation s mapping t upon v. Since u, v, t are collinear, $K(u + v) = K(u + t) = K(t + v)$, and consequently rs is a $K(u + v)$ — L-transformation mapping u upon v.

Thus we have proved that there exists to every pair of different points u, v which are both neither on K nor on L a $K(u + v)$ — L-transformation which maps u upon v; and hence it follows immediately from Theorem 3.1 that the plane is K — L-transitive.

If L is a line and p is a point not on L, then we say that the plane is

$p - L$-*homogeneous*, if there exists to every line W which is different from L and which does not pass through p and to every point v on $p + WL$ which is different from both p and WL a $p - L$-duality which maps W upon v. Note that the restrictions in the choice of v are due to the fact that $p - L$-dualities map WL upon $p + WL$ so that $WL < W$ implies $v < p + WL$.

THEOREM 3. 3. *If the point p is not on the line L, and if the plane is $p - L$-homogeneous, then the plane is $p - L$-transitive.*

Proof. Suppose that u and v are two points not on L and different from p such that u, v, p are collinear. Let W be any line different from L and $u + v$ which contains the point $L(u + v)$. There exists a $p - L$-duality $\boldsymbol{d}$ which maps W upon u and there exists a $p - L$-duality $\boldsymbol{e}$ which maps W upon v. The quotient $\boldsymbol{d}^{-1}\boldsymbol{e}$ is a $p - L$-transformation which maps u upon v so that the plane is $p - L$-transitive.

Note that, conversely, the plane is $p - L$-homogeneous, if it is $p - L$-transitive and if there exists at least one $p - L$-duality.

If K and L are two different lines, then the plane is said to be $K - L$-*homogeneous,* if it is $p - L$-homogeneous for every point p on K which is different from KL.

THEOREM 3. 4. *If the plane is $K - L$-homogeneous, then it is $K - L$-transitive.*

For let p and q be two different points on K which are both different from KL. Then it follows from Theorem 3. 3 that the plane is both $p - L$- and $q - L$-transitive; and it follows from Lemma 3. 2 that the plane is $K - L$-transitive.

4. Cartesian coördinates.[3] Suppose that X, Y, Z are three different points on a line L and that O is a point not on L. (Since these four points will play an exceptional rôle in the following discussion, it may be allowed to designate them by capitals.) We denote by F the system of those points on $O + Y$ which are different from Y (and therefore not on L). The elements in F will serve us as coördinates and we shall introduce later on addition and multiplication of the elements in F.

If c is an element in F, then $X + c$ is a line and we shall designate this line—for reasons that will be apparent very soon—by $(y = c)$.

---

[3] Since we give preference to the line L, our approach may be described as affine and the coordinates we introduce are inhomogeneous. That there are inner reasons for doing this, may be seen from the fact that the possibility of introducing homogeneous coordinates—whatever that may mean—is essentially equivalent to the validity of Desargues' Theorem—a proof of this statement will be given elsewhere.

If c is an element in F, then $Z + c$ is a line which is different from the line $O + X$. They meet in the point $(Z + c)(O + X)$ which is not on L, since Z and X are different points on L, and since neither c nor O is on L. Thus $Y + (Z + c)(O + X)$ is a well-determined line different from L which we shall designate by $(x = c)$.

Noting that $(X + c)(O + Y) = c$, it follows that $(y = c) = (y = d)$ if, and only if, $c = d$; and noting that $(O + Y)(Z + (x = c)(O + X)) = c$, it follows that $(x = c) = (x = d)$ if, and only if, $c = d$.

If c and d are elements in F, then $(x = c)$ and $(y = d)$ are different lines which meet in a point not on L, since $Y \neq X$; and we denote this point by (c, d). From $(c, d) = (x = c)(y = d)$, and from the last paragraph, it follows immediately that $(c, d) = (c', d')$ if, and only if, $c = c'$, $d = d'$.

Suppose now that p is a point not on L. Then the lines $p + X$ and $O + Y$ meet in a well-determined point not on L; and likewise $(O + Y)(Z + (p + Y)(O + X))$ is a well-determined point not on L; and one verifies that

$$p = ((O + Y)(Z + (p + Y)(O + X)), (O + Y)(p + X)).$$

Thus the points not on L are in 1 — 1-correspondence with the pairs (c, d) of elements in F.

In order to be able to introduce addition in F we have to assume that the *projective plane is Y — L-transitive.* If c is any element in F, then there exists—because of Y — L-transitivity and Theorem 2. 1—one and only one Y — L-transformation $f = f(c)$ such that $c = O^f$. If c and d are any two elements in F, then we define:

$$c + d = O^{f(c)f(d)}$$

and it is clear that an isomorphism is defined by mapping the Y — L-transformation f upon the element O^f in F so that in particular the *additive system $F = F_+$ is a group*; though thus far we have no reason to assume that F_+ be commutative.

If f is any Y — L-transformation, then $(x = c)^f = (x = c)$, since $(x = c)$ is a line through Y. If c is any element in F, then c is a point on $O + Y$, $c = O^{f(c)}$; and we find that

$$(y = c)^f = (c + X)^f = c^f + X^f = O^{f(c)f} + X = (y = c + O^f),$$

since $f = f(O^f)$. ✓

If p is any point not on L, then $p = (c, d)$ for suitable c, d in F; and we find that

$$p^f = (c, d + O^f).$$

If p is any element in F, then (O, p), (p, O) and Z are three collinear points. If q is any other element in F, then $(O, p)^{f(q)} = (O, p + q)$, $(p, O)^{f(q)} = (p, q)$ and Z are collinear points too; and this shows that the lines through Z which are different from L are characterized by the equation:

$$x + y = c \quad \text{or} \quad y = -x + c.$$

In order to be able to characterize all the lines different from L by means of equations—and at the same time to describe the points on L—we need a suitable definition of multiplication in F; and in order to do this we have to give preference to some element different from 0 which shall serve as a left-unit of this multiplication. This element may otherwise be chosen at random, though different choices lead to different definitions of multiplication.

Thus let 1 be some element in F which is different from 0. If r and s are elements in F, then $(x = r)$ and $O + (1, s)$ are two different lines which meet in a point not on L. Thus we may define the product rs as the uniquely determined element in F satisfying:

$$(r, rs) = (x = r)(O + (1, s)).$$

If M is any line which does not pass through Y — is not of the form $(x = c)$ — then M meets $O + Y$ in a point $(0, m)$ for suitable m in F and $M^{f(-m)}$ is a line passing through O. This line meets $(x = 1)$ in a point $(1, s)$ for s in F; and it is an immediate consequence of our definition of multiplication that

$$M^{f(-m)} = O + (1, s) = (y = xs)$$

so that

$$M = (O + (1, s))^{f(m)} = (y = xs + m).$$

The lines $(y = xs + m)$ and $(y = xt + n)$ meet in a point on L if, and only if, $s = t$; and thus the common point of all the lines $(y = xs + m)$ for variable m may very well be denoted by $L(s)$ — and Y as the common point of all the lines $(x = c)$ should then be denoted by $L(\infty)$. We note furthermore that not only does every line determine an equation; but the lines exhaust the admissible equations; and two lines are equal if, and only if, their equations are identical.

Since

$$(0, 0) = O = (x = 0)(O + (1, s)) = (0, 0s),$$

and since

$$(r, r0) = (x = r)(O + (1, 0)) = (x = r)(y = 0) = (r, 0),$$

it follows that $r0 = 0 = 0s$. Since

$$(1, 1s) = (x = 1)(O + (1, s)) = (1, s),$$

it follows that $1s = s$. Since

$$O + (1, -1) = (y = x(-1)) = O + Z = (y = -x),$$

it follows that $-x = x(-1)$. Since the lines $(y = xr)$ and $(y = xs + t)$ for $r \neq s$ meet in one and only one point which is not on L, it follows that for given elements r, s, t in F such that $r \neq s$, there exists one and only one element x in F satisfying $-xs + xr = t$. Since the points $(r, 0)$ and (s, t) for $r \neq s$ determine one and only one line of the type $(y = xm + n)$ where the elements m and n in F satisfy the equations $0 = rm + n$ or $n = -rm$ and $t = sm + n = sm - rm$, it follows that for given elements r, s, t in F such that $r \neq s$, there exists one and only one element x in F satisfying $sx - rx = t$.

For convenient enunciation of our results let us introduce a few terms. A *Cartesian number system* is a set F of elements with double composition: addition $m + n$ and multiplication mn, subject to the following rules.

(1) F is a group with regard to addition.

(2) The product mn of the elements m and n (in this order) is a uniquely determined element in F.

(3) $0m = m0 = 0$ for every m in F.

(4) There exists an element $1 \neq 0$ in F satisfying $1m = m$, $m(-1) = -m$ for every m in F.

(5) If r, s, t are elements in F such that $r \neq s$, then there exists one and only one element x and one and only one element y in F such that $-xs + xr = t$, $sy - ry = t$.

If F is a Cartesian number system, then denote by $P(F)$ the system of all the pairs (p, q) for p and q in F together with symbols $L(p)$ for p in F and a symbol $L(\infty) = Y$. To derive from this system $P(F)$ of points a projective plane, it suffices to state which sets of points are the sets of all the points on a given line.

a. The line L consists of all points $L(p)$ and Y.

b. The line $(x = c)$ consists of Y and of all the points (c, y).

c. The line $(y = xr + s)$ consists of $L(r)$ and of all the pairs $(x, xr + s)$.

The set $P(F)$ together with this definition of line constitutes the *projective plane over F*. That this really is a projective plane in the sense of Section 1. is readily verified.

If c is some element in F, then a $1 - 1$-correspondence f between the points of the projective plane over F is defined by

$$(x, y)^f = (x, y + c), \qquad p = p^f \text{ for } p \text{ on } L.$$

This transformation f leaves L and every line $(x = d)$ invariant and it maps the line $(y = xr + s)$ upon the line $(y = xr + s + c)$, i. e., f is an $L(\infty) - L$-transformation of the plane. This shows that the projective plane over F is $L(\infty) - L$-transitive.

Collecting all the results of this section we obtain now the following

THEOREM 4. 1.[4] *The projective plane is $p - R$-transitive for $p < R$ if, and only if, there exists a Cartesian number system F and a transformation of the projective plane upon the projective plane over F which preserves collinearity and maps p upon $L(\infty)$ and R upon L.*

5. Transitive and homogeneous planes. On account of the result of the last section we may restrict our attention to the consideration of projective planes over Cartesian number systems.

THEOREM 5. 1. *The projective plane over the Cartesian number system F is $(O + L(\infty)) - L$-transitive if, and only if, the following conditions are satisfied by F.*

(6) $$r(s + t) = rs + rt \text{ for } r, s, t \text{ in } F.$$

(7) $$r(st) = (rs)t \text{ for } r, s, t \text{ in } F.$$

Proof. If the projective plane over the Cartesian number system F is $(O + L(\infty)) - L$-transitive, then it is, in particular, $O - L$-transitive; and there exists therefore to every given element $r \neq 0$ in F an $O - L$-transformation $h = h(r)$ which maps $(1, 0)$ upon $(r, 0)$.

Since h leaves all the points on L invariant, it follows that

$$(x = 1)^h = ((1, 0) + Y)^h = (r, 0) + Y = (x = r).$$

Since every line through O is left invariant by h, and since $(1, s)$ is a point on the line $(y = xs)$ through O, it follows that

$$(1, s)^h = (x = 1)^h (y = xs)^h = (x = r)(y = xs) = (r, rs).$$

Consequently

$$(y = s)^h = ((1, s) + L(O))^h = (r, rs) + L(O) = (y = rs),$$
$$(x = s)^h = ((s, -s) + L(\infty))^h = (s, -s)^h + L(\infty)$$
$$= (y = -x)^h (y = -s)^h + L(\infty) = (y = -x)(y = r(-s)) + L($$
$$= (-r(-s), r(-s)) + L(\infty) = (x = -r(-s))$$

---

[4] It should be mentioned here that this theorem provides a convenient machinery with which to translate the formal laws of addition and multiplication—apart from statements concerning commutativity—into conditions postulating q-S-transitivity for suitable points q and lines S; and this will be done in a forthcoming publication by H. F. Gingerich.

and hence finally

$$(x, y)^h = (-r(-x), ry).$$

Considering now that (p, pt) is a point on the line $(y = xt)$ which is left invariant by $\boldsymbol{h}$, it follows that

$$(p, pt)^h = (-r(-p), r(pt)) = (-r(-p), (-r(-p))t)$$

so that

$$r(pt) = (-r(-p))t \text{ for } r, p, t \text{ in } F.$$

Furthermore we have

$$(y = xs + t)^h = (L(s) + (0, t))^h = L(s) + (0, rt) = (y = xs + rt)$$

and since $(1, s + t) < (y = xs + t)$, it follows that $(r, r(s + t)) < (y = xs + rt)$ so that $r(s + t) = rs + rt$, verifying (6). This left-distributive law implies in particular $r(-s) = -rs$ so that we have

$$r(pt) = (-r(-p))t = (rp)t, \text{ i. e., the associative law (7).}$$

If, conversely, (6) and (7) are satisfied by F, then the transformation $\boldsymbol{h}$ defined by

$$(x, y)^h = (rx, ry) \text{ for some } r \neq 0$$

is one-to-one and maps the line $(x = c)$ upon the line $(x = rc)$ and the line $(y = xs + t)$ upon the line which consists of the points $(rx, r(xs + t)) = (rx, r(xs) + rt) = (rx, (rx)s + rt)$, i. e., upon the line $(y = xs + rt)$ so that $\boldsymbol{h}$ clearly induces an $O - L$-transformation mapping $(1, 0)$ upon $(r, 0)$; and this shows—Theorem 3. 1—the $O - L$-transitivity of the plane. Applying Lemma 3. 2 and Theorem 4. 1 it follows that the plane is $(O + L(\infty)) - L$-transitive.

THEOREM 5. 2. *The projective plane over the Cartesian number system F is $(O + L(\infty)) - L$-homogeneous if, and only if, F is an ordinary commutative field.*

Proof. If the projective plane over F is $(O + L(\infty)) - L$-homogeneous, then it certainly is—by Theorem 3. 4—$(O + L(\infty)) - L$-transitive. It follows, therefore, from Theorem 5. 1 that F is a group with regard to multiplication (excepting 0) and that F is left-distributive. Thus all we have to prove is the commutativity of multiplication.

The projective plane is in particular $O - L$-homogeneous. Hence there exists an $O - L$-duality $\boldsymbol{d}$ of the projective plane which maps the line $(x = 1)$ upon the point $(0, -1)$. (This choice is admissible, since $(x = 1) \neq L$, and since $(0, -1) < (x = 0) = O + L(\infty) = O + L(x = 1)$.)

We note that $\boldsymbol{d}$ interchanges $(y = xa)$ and $L(a)$, $(x = 0)$ and $L(\infty)$, L and O. Hence we find that

$$\begin{aligned}(1, a)^d &= ((x = 1)(y = xa))^d = (x = 1)^d + (y = xa)^d \\ &= (0, -1) + L(a) = (y = xa - 1), \\ (y = c)^d &= ((1, c) + L(0))^d = (y = xc - 1)(y = 0) \\ &= (c^{-1}, 0) \quad \text{for} \quad c \neq 0,\end{aligned}$$

since F is a group with regard to multiplication;

$$\begin{aligned}(c, c)^d &= ((y = x1)(y = c))^d = L(1) + (c^{-1}, 0) \\ &= (y = x - c^{-1}) \quad \text{for} \quad c \neq 0\end{aligned}$$

(considering $r1 = 1r = r$),

$$\begin{aligned}(x = c)^d &= ((c, c) + L(\infty))^d = (y = x - c^{-1})(x = 0) \\ &= (0, -c^{-1}) \quad \text{for} \quad c \neq 0, \\ (p, q)^d &= ((x = p)(y = q))^d = (0, -p^{-1}) + (q^{-1}, 0) \\ &= (y = xqp^{-1} - p^{-1})\end{aligned}$$

for q and p different from 0, again using the fact that F is a group under multiplication.

Suppose now that r and s are two elements in F. If one of them is 0, then certainly $rs = sr$. If none of them is 0, then we consider the point $(r, rs) < (y = xs)$. Since $\boldsymbol{d}$ is a duality, we find that

$$L(s) = (y = xs)^d < (r, rs)^d = (y = x\,rs\,r^{-1} - r^{-1})$$

and this implies that

$$(y = x\,rsr^{-1} - r^{-1}) = (y = xs - r^{-1}).$$

Consequently $s = rsr^{-1}$ or $sr = rs$ as was to be proved.

Assume now, conversely, that F is an ordinary commutative field. Then it follows from Theorem 5.1 that the projective plane over F is $(O + L(\infty)) - L$-transitive. Hence it suffices to prove the existence of one $O - L$-duality in order to assure $(O + L(\infty)) - L$-homogeneity.

A duality $\boldsymbol{d}$ which meets our requirements may be defined as follows.

$$\begin{aligned}O^d = L, \quad & (0, c)^d = (x = c^{-1}) \quad \text{for} \quad c \neq 0, \\ & (c, d)^d = (y = xdc^{-1} - c^{-1}) \quad \text{for} \quad c \neq 0, \\ L(\infty)^d = {} & (x = 0), \quad L(a)^d = (y = xa); \\ L^d = O, \quad & (x = 0)^d = L(\infty), \quad (y = xa)^d = L(a), \\ & (x = c)^d = (0, -c^{-1}) \quad \text{for} \quad c \neq 0, \\ & (y = xa + c)^d = (c^{-1}, c^{-1}a) \quad \text{for} \quad c \neq 0.\end{aligned}$$

Combining this last theorem with the well-known result that a projective plane is a projective plane over an ordinary commutative field if, and only if, the theorems of Desargues and Pappus are valid in this plane, we find:

The theorems of Desargues and Pappus are valid in a projective plane if, and only if, there exist two different lines K and L such that the plane is $K-L$-homogeneous.

6. The characteristic configurations. Suppose that p is a point not on the line L. We intend to characterize $p-L$-homogeneity by means of a certain configuration. With this in mind we introduce the following

$p-L$-property: Whenever $x(i)$, $y(i)$, $z(i)$ for $i=1,2$ are points such that

$$y(1)\neq y(2),\quad z(1)\neq z(2),\quad y(i)\neq z(i),$$
$$y(i)<L,\quad z(i)<L,\quad x(i) \text{ not on } L, \text{ and}$$
$$L(x(2)+p),\quad (x(1)+y(1))(p+y(2)),\quad (x(1)+z(1))(p+z(2))$$

are three different collinear points,[5] *then*

$$L(x(1)+p),\quad (x(2)+y(2))(p+y(1)),\quad (x(2)+z(2))(p+z(1))$$

are three different collinear points.

The second of these triplets of collinear points is obtained from the first one by interchanging 1 and 2.[6] Thus the $p-L$-property implies the following fact:

$$L(x(2)+p),\quad (x(1)+y(1))(p+y(2)),\quad (x(1)+z(1))(p+z(2))$$

are three different collinear points if, and only if,

$$L(x(1)+p),\quad (x(2)+y(2))(p+y(1)),\quad (x(2)+z(2))(p+z(1))$$

are three different collinear points.

THEOREM 6.1. *The projective plane is $p-L$-homogeneous ($p \nless L$) if, and only if, the $p-L$-property is satisfied by the points of the plane.*

Proof. Assume first that the projective plane is $p-L$-homogeneous, that $x(i)$, $y(i)$, $z(i)$ are, for $i=1,2$, points such that

$$y(1)\neq y(2),\quad z(1)\neq z(2),\quad y(i)\neq z(i),\quad y(i)<L,\quad z(i)<L,$$

---

[5] Note e.g. that this implies $x(2)\neq p$, since otherwise $L(x(2)+p)=0$ and not a point.

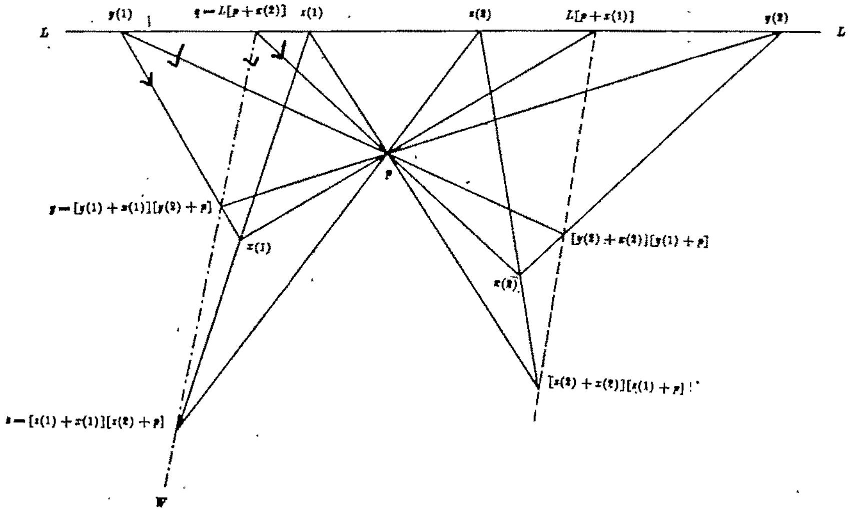
[6] H. Liebmann has introduced two exchange postulates which are rather different from the $p-L$-property though his "great exchange postulate" seems to be equivalent to the condition that the $p-L$-condition should be satisfied for every p not on L which is on a given line $K\neq L$; cf. M. Steck: *Deutsche Mathematik* 1 (1936), 165-174.

$x(i)$ not on L, and such that

$$q = L(x(2) + p), \quad y = (x(1) + y(1))(p + y(2))$$

and $z = (x(1) + z(1))(p + z(2))$ are three different collinear points. Put $W = q + y + z$.

If W and L were equal, then y would be on L, and since $x(1)$ and p are not on L, it would follow that $y = y(1) = y(2)$ which is impossible. If W and $p + q$ were equal, then p, y and z would be equal points which contradicts our hypothesis that $y \neq z$. Since the three lines W, L, $p + x(2)$ meet in the point q, and since the plane is p — L-homogeneous, there exists a p — L-duality $\boldsymbol{d}$ which maps W upon $x(2)$.



Since p is not on W, $y = W(y + p)$. Since $y + p = p + y(2)$, it follows that

$$y^d = x(2) + y(2) \quad \text{and likewise} \quad z^d = x(2) + z(2).$$

Since

$$y + x(1) = x(1) + y(1) = y + y(1),$$

it follows that

$$(y + x(1))^d = (x(2) + y(2))(p + y(1))$$

and likewise that

$$(z + x(1))^d = (z(2) + x(2))(p + z(1))$$

and finally we have

$$(p+x(1))^d = L(p+x(1)).$$

Since the three lines on the left are three different lines passing through the point $x(1)$—otherwise either $z(1)=z(2)$ or $y(1)=y(2)$ or $y(1)=z(1)$—it follows that the three expressions on the right signify three different collinear points; and this proves that the $p - L$-property is a consequence of $p - L$-homogeneity.

Assume now, conversely, that W is a line different from L, that p is not on L and that $x(2)$ is a point on $p+WL$ which is different from both p and $WL=q$. We define a function $\boldsymbol{d}$ of the points of our plane as follows.

$$p^d = L,\ r^d = p + r \text{ for } r < L,$$
$$v^d = x(2) + (p+v)L \text{ for } v < W.$$

If finally $x(1)$ is a point neither on W nor on L, then let y be some point on W which is neither on L nor on $p+x(1)$. We put

$$\begin{aligned} d(x(1),y) &= y^d(L(x(1)+y))^d + (p+x(1))^d \\ &= (x(2)+(p+y)L)(p+(x(1)+y)L) + (p+x(1))L \\ &= (x(2)+y(2))(p+y(1)) + (p+x(1))L \end{aligned}$$

and it is a consequence of the $p - L$-property that $d(x(1),y) = d(x(1),z)$, if z is another point on W meeting all the requirements. Thus we may put finally $x(1)^d = d(x(1),y)$ for y any point on W, not on L or $p+x(1)$. It is an immediate consequence of our construction that $\boldsymbol{d}$ is a $1-1$-correspondence mapping points upon lines and collinear points upon collinear lines; and thus $\boldsymbol{d}$ defines a duality which meets all the requirements.

THEOREM 6.2. *The projective plane is $p - L$-transitive (where the point p may or may not be on the line L) if, and only if, the following condition is satisfied.*

Whenever $b(i)$, $c(i)$ for $i = 1, 2, 3$ are triplets of points such that

$$b(i),\ c(i),\ p \text{ are collinear for } i = 1, 2, 3,$$

$(b(i)+b(j))(c(i)+c(j))$ is for $i \neq j$ a point different from p and such that at least two different ones of the points $(b(i)+b(j))(c(i)+c(j))$ for $i \neq j$ are on L, then the third one is on L too.

The condition of this theorem may be described as the $p - L$-Desargues condition.—The customary enunciation of Desargues' Theorem:

Two triangles are perspective from some point if, and only if, corresponding sides meet in collinear points, though sounding more elegant is quite unsatisfactory, since there do not exist projective planes with this property.

Proof. Suppose first that the plane is $p - L$-transitive, that $b(i)$, $c(i)$ are points such that $b(i)$, $c(i)$, p are collinear for $i = 1, 2, 3$, $d(i, j) = (b(i) + b(j))(c(i) + c(j))$ for $i \neq j$ are points different from p and that at least two of these points, say $d(1, 2)$ and $d(1, 3)$, are on L. Since $b(1)$, $c(1)$, p are three different collinear points, and since neither $b(1)$ nor $c(1)$ is on L — since otherwise $c(1, 2) = c(1, 3)$ — there exists a $p - L$-transformation f mapping $b(1)$ upon $c(1)$. If $i \neq 1$, then

$$\begin{aligned} b(i)^f &= (b(i) + p)^f (b(i) + c(i, 1))^f = (b(i) + p)(b(1) + c(i, 1))^f \\ &= (c(i) + p)(c(1) + c(i, 1)) = (c(i) + p)(c(1) + c(i)) = c(i) \end{aligned}$$

so that $L(b(i) + b(j)) = L(c(i) + c(j)) = d(i, j)$ as was to be shown.

Assume, conversely, that the condition of the theorem holds true; and that r, s are points not on L and different from p such that p, r, s are collinear. Let q be some point on L such that p, r, s, q are not collinear. Then the lines $R = r + q$ and $S = s + q$ are different from L, pass through q, but not through p; and it will clearly be sufficient to prove the existence of a $p - L$-transformation mapping R upon S.

If x is any point of the plane, not on L, then let k be some point on L which is different from $L(p + x)$; and put

$$f(x; k) = (p + x)(k + S(p + (x + k)R)).$$

It is then an immediate consequence of the $p - L$-Desargues condition, that

$$f(x; k) = f(x; h) \text{ for } k \text{ and } h \text{ points on } L, \text{ not on } p + x;$$

and a transformation f meeting all the requirements may therefore be defined by the formulas:

$$\begin{gathered} x = x^f \text{ for } x \text{ on } L,\ p = p^f, \\ x^f = f(x; k) \text{ for } x \text{ not on } L,\ x \neq p, \end{gathered}$$

and k any point on L, not on $x + p$.

UNIVERSITY OF ILLINOIS,
URBANA, ILLINOIS.

SQUARE ROOTS OF MEASURE PRESERVING TRANSFORMATIONS.*

By PAUL R. HALMOS.

1. **Introduction.** Let X be a metric, complete, separable space, and let $m(E)$ be a measure defined on a Borel field of subsets of X. Sets of this Borel field will be called measurable sets. Concerning m we assume that spheres are measurable and have positive measure, $m(X) < \infty$, and that every measurable set is contained in a G_δ of the same measure.[1] Let T be a one to one, measure preserving, (not necessarily continuous), transformation of X into itself. A classical problem concerning such measure preserving transformations is their embeddability into a flow: does there exist a one parameter group of measure preserving transformations T_t, satisfying suitable regularity requirements with respect to the real parameter t, and also the group condition $T_tT_s = T_{t+s}$, for which $T_1 = T$? Other problems of interest center around the subject of change of velocities in a flow: what purely geometric properties of a flow remain invariant when we replace T_t by another flow S_s related to it by the sole requirement that the complete trajectory of every point remains the same? One discrete analog of this latter problem might be the replacement of a transformation T by another transformation S related to it by the requirement that the orbit under T of every point is a subset of its orbit under S. It is the purpose of this paper to make a first step in both these directions by finding necessary and sufficient conditions for the existence of a transformation S, called a square root of T, for which $S^2 = T$.[2] Clearly the existence of a square root is a necessary condition for embeddability into a flow, and the finding of a square root is the simplest speeding up we can effect on a transformation.

Consider the unitary operator U defined in $L_2(X)$ by $Uf(x) = f(Tx)$. The correspondence between the set of all measure preserving transformations

---

* Received September 17, 1940; Revised May 25, 1941.

[1] In the terminology of J. von Neumann, ("Einige Sätze über messbare Abbildungen," *Annals of Mathematics*, vol. 33 (1932), pp. 574-586), X is an m-space and the measure is obtained from an 1-measure. We shall not make explicit use of the topological properties of X, but we assume them in order to be able to make use of von Neumann's results. See also remark (C) in § 4.

[2] The equality of the two transformations, S^2 and T, as well as every statement we make concerning the vanishing of a function, the existence of a function or of a transformation, the equality of two sets or two functions, etc., is to be qualified, throughout this paper, by the phrase "almost everywhere."

T and those unitary operators U which are obtained from them is one to one, and the analytical tools available in the study of operators make it easier to investigate U. We shall give a complete solution of our problem in case U has pure point spectrum: in other words, in case there exists a sequence of complex numbers λ_n, (characteristic values of U), and a complete orthogonal set of functions $f_n(x)$, (characteristic functions of U), such that $Uf_n = \lambda_n f_n$. The principal tool that we shall use is the following result, due to von Neumann, which, for convenience of reference, we state in the form of a lemma.

LEMMA 1. *If U is a unitary operator on $L_2(X)$ such that for every pair f, g of bounded measurable functions*

$$(*) \qquad U(fg) = Uf\,Ug$$

then there exists on X a one to one measure preserving transformation T such that $Uf(x) = f(Tx)$.[3]

In order to facilitate the application of this lemma we make the following remarks.

(A) If U satisfies $(*)$ for f, g in a complete set in L_2 then U satisfies $(*)$ for all f, g. This is easily verified by letting f and g be linear combinations of functions in the complete set and limits of such linear combinations.

(B) If M is the set of functions in L_2 satisfying $Uf = \lambda f$ for some fixed λ then there exists in M a complete orthogonal set of bounded functions. For, let $f_n(x)$, $n = 1, 2, \cdots$, be a complete set in M. For every positive integer k let $f_n^k(x) = f_n(x)$ whenever $|f_n(x)| \leq k$; $f_n^k(x) = 0$ otherwise. Then $f_n^k(x) \epsilon M$ and together the $f_n^k(x)$ are a complete set in M; by an application of the Schmidt orthogonalization process[4] we obtain the desired set.

(C) If E is a measurable subset of X and if U' is a unitary operator on $L_2(E)$ satisfying $(*)$ for all bounded f, g in $L_2(E)$, then there exists on E a one to one measure preserving transformation T' such that $U'f(x) = f(T'x)$. For, we may apply Lemma 1 to the operator U defined by the relations $Uf = U'f$ if $f = 0$ on the complement of E, $Uf = f$ if $f = 0$ on E.

We call attention to the following notations, used throughout the paper. If E is any subset of X, $\bar{E}$ will denote the complementary set $X - E$. If P is any proposition concerning points $x \epsilon X$, ${}_x\{P\}$ will denote the set of points where P is true. T will always be a given one to one and measure preserving transformation of X into itself, and U its corresponding unitary operator.

---

[3] Here fg is the ordinary product $f(x)g(x)$. See J. von Neumann, "Zur Operatorenmethode in der Klassischen Mechanik," *Annals of Mathematics*, vol. 33 (1932), pp. 587-642, particularly pp. 618-619.

[4] See R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Springer, 1931, vol. 1, p. 41.

2. The necessary and sufficient conditions. We are now in a position to begin the characterization of transformations which have square roots.

THEOREM 1. *If T has a square root, $T = S^2$, then -1 is a characteristic value of U of even multiplicity* $(= 0, 2, 4, \cdots, \infty)$.

Proof. We shall prove the theorem by making correspond to each characteristic function $f(x)$ belonging to the characteristic value -1 another such characteristic function $f^*(x)$ with the following properties: (i) $f(x)$ and $f^*(x)$ are orthogonal; (ii) if $f_2(x)$ is orthogonal to $f_1(x)$ and $f^*_1(x)$ then so is $f^*_2(x)$; (iii) $f^{**}(x) = -f(x)$.

Suppose that $f(Tx) = -f(x)$. We define $f^*(x) = \overline{f(Sx)}$.[5] We have

$$f^*(Tx) = \overline{f(STx)} = \overline{f(TSx)} = -\overline{f(Sx)} = -f^*(x);$$

it remains to be proved that $f^*(x)$ has the properties (i), (ii), (iii).

Since S is a measure preserving transformation we have, for any integrable function $u(x)$, $\int u(x)\,dx = \int u(Sx)\,dx$. Writing $u(x) = f(x)\overline{f^*(x)}$, it follows easily that $\int u(x)\,dx = -\int u(Sx)\,dx = 0$. Next: if $f_1(x)$ and $f_2(x)$ are orthogonal characteristic functions of T belonging to the characteristic value -1, and if $f_2(x)$ is also orthogonal to $f^*_1(x) = \overline{f_1(Sx)}$, then an application of S to the arguments of the integrands in $\int f_1(x)\overline{f^*_2(x)}\,dx$ and $\int f^*_1(x)\overline{f^*_2(x)}\,dx$ shows that $f^*_2(x)$ is orthogonal to both $f_1(x)$ and $f_2(x)$. Finally it follows from the definition of $f^*(x)$ that $f^{**}(x) = -f(x)$. This completes the proof.

The result of Theorem 1 motivates the following theorem which will be used to split up the study of transformations, satisfying the necessary condition of Theorem 1, into two cases. In Theorem 2 and in much of what follows we shall make repeated use of a concept for which it is therefore desirable to introduce a convenient notation. If λ is a characteristic value of U and E is a subset of X, we shall denote by (E, λ) the class of characteristic functions belonging to the characteristic value λ, which vanish on $\bar{E}$. We shall refer to a member of this class as an (E, λ)-function.

THEOREM 2. *To every T there corresponds a measurable set E, invariant under T, such that there are no $(E, -1)$-functions other than $f(x) \equiv 0$, and such that there is at least one $(\bar{E}, -1)$-function which does not vanish on $\bar{E}$.*[6]

---

[5] $\bar{z}$ is the conjugate of the complex number z.

[6] E may be the whole space or the empty set. In the case where E turns out to be X the theorem says that -1 is not a proper eigenvalue of U, (i. e. the only charac-

Proof. Let M be the set of functions $f(x)$ in L_2 for which $f(Tx) = -f(x)$. If M is of dimension 0 we write $E = X$, (see footnote (6)); otherwise let $f_1, f_2, \cdots$ be a complete orthogonal set in M. Since along with f_n, $R(f_n)$ and $I(f_n)$ also belong to M,[7] and since the set

$$\{R(f_n), I(f_n);\quad n = 1, 2, \cdots\}$$

is complete, the Schmidt orthogonalization process permits us to assume that the f_n are real valued. Let E be the set

$$E = {}_x\{f_1 = 0\} \cdot {}_x\{f_2 = 0\} \cdot \cdots.$$

Since $f_n(Tx) = -f_n(x)$ the set ${}_x\{f_n = 0\}$ is invariant under T, and therefore so is E. Suppose now that $f(x)$ is a bounded $(E, -1)$-function. Since we may write $f(x) = \sum_{i=1}^{\infty} a_i f_i(x)$, (the Fourier series converging in the mean), we have $f(x) = 0$ on E; since by hypothesis $f(x) = 0$ on $\tilde{E}$, it follows that $f(x) \equiv 0$.

To construct an $(\tilde{E}, -1)$-function which does not vanish on $\tilde{E}$ we proceed as follows. Let A_n be the set ${}_x\{f_1 = 0\} \cdot \cdots {}_x\{f_{n-1} = 0\}$, $n = 2, 3, \cdots$; $A_1 = X$, and let $a_n(x)$ be the characteristic function of A_n. Write $f'(x) = \sum_{n=1}^{\infty} a_n(x) f_n(x)$. The series converges everywhere since, as is easily verified, for each x at most one term is different from zero. We may now define

$$f(x) = \frac{f'(x)}{1 + |f'(x)|},$$

then $f(x)$ is an $(\tilde{E}, -1)$-function in L_2, (in fact $f(x)$ is bounded), and $f(x)$ does not vanish on $\tilde{E}$; we remark that $f(x)$ is real valued by construction.

Retaining the notation of the above proof we have

LEMMA 2. *If T has a square root S, then $SE = E$.*

Proof. If $E = X$ or $E = 0$, the lemma is trivial. In any other case let $f(x)$ be an $(\tilde{E}, -1)$-function which does not vanish on $\tilde{E}$; write $G = {}_x\{g(x) = 0\}$, where $g(x) = f(S^{-1}x)$. Since $E = {}_x\{f(x) = 0\}$, we have $SE = G$. Since, furthermore, $g(Tx) = -g(x)$, it follows that $g(x) = 0$ on E, so that $E \subseteq G$; since, (S being a measure preserving transformation), $m(E) = m(G)$, we have $E = G = SE$.

---

teristic function corresponding to -1 is $f(x) \equiv 0$), and the condition on $\tilde{E}$ is satisfied vacuously. In the case where $E = 0$, the theorem says that there exists a characteristic function corresponding to -1 which never vanishes.

[7] $R(z)$ and $I(z)$ are the real and imaginary parts, respectively, of the complex number z.

Considering the transformation T only on the subspaces E and $\tilde{E}$ we observe that the last two results enable us to split the problem of finding square roots into two cases.

Case I. The only function satisfying $f(Tx) = -f(x)$ is $f(x) \equiv 0$.

Case II. There exists a nowhere vanishing real valued function $f(x)$ for which $f(Tx) = -f(x)$.

The only detail we have to verify is whether or not Lemma 1 is applicable to the subspaces E and $\tilde{E}$. This is answered in the affirmative by the remark (C) following Lemma 1.

We shall be able to prove that if T has a pure point spectrum[8] and belongs to case I then T always has a square root. First, however, we shall prove a special case of this theorem because the special case is sufficiently important to make a simple proof desirable and because the proof of the special case will serve as a guiding thread in the steps of the general proof. We recall a definition: T is said to be metrically transitive if T has no non-trivial invariant sets: more precisely, if $TE = E$ implies $m(E)m(\tilde{E}) = 0$.

THEOREM 3. *If T is a metrically transitive transformation with a pure point spectrum not containing -1 then T has a square root.*

The proof of this theorem depends on the following lemma.

LEMMA 3. *If G is a countable multiplicative group of complex numbers of modulus one which does not contain -1, then there exists an isomorphic group G' such that if λ and λ' correspond under the isomorphism then $(\lambda')^2 = \lambda$.*

Proof of Lemma 3. We observe that it is impossible that both the numbers λ and $-\lambda$ should be in G, for any λ, for then we should have $(\lambda)(-1/\lambda) = -1 \epsilon G$, contrary to our hypothesis. Hence if we denote by H the subgroup of squares in G, then to each $\mu \epsilon H$ there corresponds a unique $\lambda \epsilon G$ for which $\lambda^2 = \mu$: throughout this proof we shall denote λ by $\sqrt{\mu}$.

Let λ_1 be an arbitrary element of H; choose λ_2 in G but not in H, λ_3 in G but not in H nor in $\lambda_2 H$ (where λH means the set of all numbers of the form $\lambda\mu$, $\mu \epsilon H$), and so on. In this way we obtain a finite or infinite sequence of elements $\lambda_1, \lambda_2, \cdots$ in G, such that the cosets $\lambda_1 H, \lambda_2 H, \cdots$ are disjoint and exhaust G. If we define $\lambda_i H \odot \lambda_j H = \lambda_i \lambda_j H$, then under this composition the cosets $\lambda_n H$ form a group isomorphic with the quotient group G/H. We observe that every element in this (abelian) quotient group is of order two, since $\lambda_n^2 \epsilon H$ implies $\lambda_n^2 H = H$ for $n = 1, 2, \cdots$. Hence we may find in this quotient group a basis represented by, say, the cosets $\lambda_{a_1}H, \lambda_{a_2}H, \cdots$. This implies that every element of G is uniquely expressible in the form

---

[8] This means, of course, that U has pure point spectrum.

$$\lambda_{a_1}^{\epsilon_1}\lambda_{a_2}^{\epsilon_2}\cdots\mu,$$

where $\epsilon_j = 0$ or 1, $\mu \epsilon H$, and only a finite number of the ϵ_j are different from zero. We now choose, for each j, either of the two values of $\sqrt{\lambda_{a_j}}$: we denote this number by ν_{a_j}. To the general element of G written above we make correspond the number

$$\nu_{a_1}^{\epsilon_1}\nu_{a_2}^{\epsilon_2}\cdots\sqrt{\mu}.$$

It is easily verified that the set of all these numbers forms a group G' with the desired properties.

Proof of Theorem 3. Since T is metrically transitive, U has simple spectrum, so that to each characteristic value λ_i there corresponds a characteristic function f_i, uniquely determined except for a constant multiplicative factor, and by hypothesis the f_i form a complete orthogonal set in L_2. Since $U(f_if_j) = (Uf_i)(Uf_j) = \lambda_i\lambda_j f_if_j$, we must have $\lambda_i\lambda_j = \lambda_k$, $f_if_j = cf_k$ for some positive integer k and constant c. The set of all characteristic values forms therefore a group G to which Lemma 3 may be applied to obtain an isomorphic square root group G'. A unique unitary operator V may be defined by the conditions $Vf_i = \lambda'_i f_i$; where λ'_i is the isomorph of λ_i in G'. It is clear that $V^2 = U$; and since we have $V(f_if_j) = cVf_k = c\lambda'_k f_k$, and $(Vf_i)(Vf_j) = \lambda'_i\lambda'_j f_if_j = \lambda'_k cf_k$, so that $V(f_if_j) = (Vf_i)(Vf_j)$, Lemma 1 and the remark (A) following it yield the proof of the theorem.

By means of this theorem it is easy to give a complete spectral characterization of the metrically transitive transformations with pure point spectrum which have a square root. In fact we can show that the sufficient condition of Theorem 3 is necessary, even without the assumption of pure point spectrum.

THEOREM 4. *If* -1 *is a characteristic value of the metrically transitive transformation* T *then* T *does not have a square root.*

This theorem is an immediate consequence of Theorem 1 and the fact that the characteristic values of a metrically transitive transformation are simple. We elect, however, to give another proof to which we shall have occasion to refer later. (See the proof of Theorem 6.)

Proof of Theorem 4. Suppose that $f(x)$ is a characteristic function belonging to the characteristic value -1. Then $f(x)$ is an invariant function under U^2, so that T^2 is not metrically transitive. Let C_0 be a non-trivial invariant set of T^2 (i. e., $0 < m(C_0) < m(X)$). Then since $C_0 + TC_0$ is invariant under T we must have $C_0 + TC_0 = X$, and since $C_0 \cdot TC_0$ is invariant under T, $C_0 \cdot TC_0 = 0$. We write $TC_0 = C_1$.

Suppose further that T has a square root S. Writing $C_{ij} = C_i \cdot SC_j$, we have $SC_{ij} = SC_i \cdot TC_j = C_{j+1} \cdot SC_i = C_{j+1,i}$ (where the subscripts are to be interpreted modulo 2), so that $TC_{ij} = C_{i+1,j+1}$. It follows that $A = C_{00} + C_{11}$, and $B = C_{01} + C_{10}$, are non-trivial invariant sets under T, contradicting the hypothesis of metrical transitivity. This proves the impossibility of the existence of S.

We return now to the more general problem, and give the proof of the following theorem.

THEOREM 5. *If T has pure point spectrum not containing -1 then T has a square root.*

Proof. For any complex number λ of absolute value one we define a set $M(\lambda)$ as follows. If λ is not a characteristic value of U we define $M(\lambda)$ to be the null set. If λ is a characteristic value, let $f_1, f_2, \cdots$ be a complete orthogonal set of functions in the manifold of characteristic functions belonging to λ. Let $K = K(\lambda)$ be the set ${}_x\{f_1 = 0\} \cdot {}_x\{f_2 = 0\} \cdot \cdots$; write $M(\lambda) = \tilde{K}$. Just as in the proof of Theorem 2 we see that there are no (K, λ)-functions, and that there is at least one $(M(\lambda), \lambda)$-function which does not vanish on $M(\lambda)$. We introduce the convention that $\sqrt{\lambda}$, for any λ of modulus one, is that value of the square root of λ whose argument lies in the interval $0 \leqq \theta < \pi$.

Consider, for any λ, the sets $M(\lambda)$ and $M(-\lambda)$. We remark that these two sets are disjoint, for otherwise we should have bounded characteristic functions, f belonging to λ, and g belonging to $-\lambda$, such that the set where neither vanishes is of positive measure. Then the function $f \cdot \bar{g}$ does not vanish identically, and it is a characteristic function belonging to $(\lambda)(\overline{-\lambda}) = -1$.

Now let λ_n be any one of the characteristic values of U; write

$$F_n = M(\sqrt{\lambda_n}), \quad G_n = M(-\sqrt{\lambda_n}), \quad H_n = M(\lambda_n) - (F_n + G_n),^{9}$$
$$K_n = X - (F_n + G_n + H_n).$$

In view of the remark above the sets F_n, G_n, H_n, K_n, for any fixed n, are disjoint and cover X.

We shall now define a unitary operator V which will turn out to be the unitary correspondent of one of the square roots of T. Because of the

---

[9] It is easy to see that $M(\lambda_n) \supseteq M(\pm\sqrt{\lambda_n})$.

hypothesis of pure point spectrum it is sufficient to define V for a complete subset of the set of all characteristic functions of U: it will then have a uniquely determined extension to all L_2. We shall explicitly define V for all (F_n, λ_n)-functions, all (G_n, λ_n)-functions, and, for each fixed n, for all $(H_n E_k, \lambda_n)$-functions, where the E_k will be a suitably chosen finite collection of disjoint measurable sets covering X. The totality of all functions here listed is obviously complete.

For every (F_n, λ_n)-function f we write $Vf = \sqrt{\lambda_n} f$; for every (G_n, λ_n)-function g, $Vg = -\sqrt{\lambda_n} g$. For (H_n, λ_n)-functions we shall give the definition of V by induction on n. For every (H_1, λ_1)-function h we define $Vh = \sqrt{\lambda_1} h$. In defining V for (H_2, λ_2)-functions we distinguish between two possibilities. (i) No odd power of λ_2 is in the (cyclic) group generated by λ_1. In this case we define, for every (H_2, λ_2)-function h, $Vh = \sqrt{\lambda_2} h$. (ii) There is an odd power of λ_2 in the group generated by λ_1. In this case we define V separately for the four possible classes of functions

$$(H_2 F_1, \lambda_2), \quad (H_2 G_1, \lambda_2), \quad (H_2 H_1, \lambda_2), \quad (H_2 K_1, \lambda_2).$$

For any function k of the last class we define $Vk = \sqrt{\lambda_2} k$. The definitions of V for the other three classes are similar to each other: we give the definition only for the first class $(H_2 F_1, \lambda_2)$. By hypothesis $\lambda_2^{n_2} = \lambda_1^{n_1}$ with n_2 odd. Then λ_2 is one of the values of $\lambda_1^{n_1/n_2}$. To the class of functions (F_1, λ_1) we have already made correspond one of the values, say ν_1, of the square root of λ_1. We now make correspond to the class $(H_2 F_1, \lambda_2)$ that uniquely determined value, ν_2, of $\nu_1^{n_1/n_2}$ which is a square root of λ_2. The existence of such a ν_2 is obvious and its uniqueness follows from the fact that n_2 is odd, so that no n_2-th root of unity is -1. It still remains to be shown that if we also have $\lambda_2^{n'_2} = \lambda_1^{n'_1}$, where n'_2 is odd, then the square root ν'_2 obtained from this relation by our process is the same as ν_2. This is true because the square root ν''_2 obtained from the relation $\lambda_2^{n_2 n'_2} = \lambda_1^{n_1 n'_1}$ is seen to be equal to both ν_2 and ν'_2. Our notation being consistent we may define Vf, for any $(H_2 F_1, \lambda_2)$-function f, by $Vf = \nu_2 f$.

The details of the general step in the inductive process are cumbersome: we merely sketch them. Suppose then that V has been defined for all characteristic functions belonging to λ_j, $j = 1, \cdots, n$; we are to define V for the class (H_{n+1}, λ_{n+1}). We do this by considering separately the classes $(H_{n+1} A_1 A_2 \cdots A_n, \lambda_{n+1})$ where each A_j may be F_j, G_j, H_j, K_j, $j = 1, \cdots, n$. In any one of these classes a certain number of A's may be K's. We consider only those numbers of the set $\lambda_1, \cdots, \lambda_n$ whose subscripts correspond to the other A's: those which are F's, G's, or H's. We distinguish between two possibilities according as there is or is not an odd power of λ_{n+1} in the group generated by the λ's under consideration. In the latter case we associate with the function class under consideration an arbitrary one of the two values of $\sqrt{\lambda_{n+1}}$; in the former case we prove as above the existence of a unique square root of λ_{n+1} compatible with the previously chosen square roots.

V is now defined; we shall show that it satisfies (*), (see Lemma 1), thereby concluding the proof of the theorem.

For each positive integer n we suppose that we have given a complete set of bounded characteristic functions belonging to λ_n such that every member of the set belongs to one of the function classes $(A_1A_2\cdots A_n, \lambda_n)$ where $A_j = F_j, G_j, H_j, K_j$, $j = 1, \cdots, n$. Let ϕ_n and ϕ_m be such characteristic functions corresponding to λ_n and λ_m respectively. If $\phi_n\phi_m \equiv 0$ then $V(\phi_n\phi_m) = 0 = (V\phi_n)(V\phi_m)$, so that (*) is satisfied in this case. If $\phi_n\phi_m \not\equiv 0$ then this product is a characteristic function belonging to the characteristic value λ_k, say, where $\lambda_k = \lambda_n\lambda_m$. In this case the function ϕ_j, $j = n, m, k$, can be written as a sum of functions $f_j, g_j, h_j^1, h_j^2, \cdots, h_j^p$, where f_j is an (F_k, λ_j)-function, g_j is a (G_k, λ_j)-function, and h_j^i is a function in the i-th class of the classes $(H_kA_1A_2\cdots A_{k-1}, \lambda_j)$ (in some order), where $A_r = F_r, G_r, H_r, K_r$; $r = 1, 2, \cdots, k-1$.

We have

$$\phi_j = f_j + g_j + h_j^1 + \cdots + h_j^p, \qquad (j = n, m, k).$$

We are to show that

$$Vf_k = Vf_n \cdot Vf_m, \quad Vg_k = Vg_n \cdot Vg_m, \quad Vh_k^i = Vh_n^i \cdot Vh_m^i, \quad (i = 1, \cdots, p).$$

Now $Vf_k = \nu_k f_k$, where ν_k is one of the square roots of λ_k. On the other hand $Vf_n \cdot Vf_m = \nu_n\nu_m f_n f_m = \nu_n\nu_m f_k$. We must have $\nu_n\nu_m = \nu_k$ for $\nu_n\nu_m$ is one of the square roots of λ_k and if it were not equal to ν_k then F_k would intersect both the sets $M(\sqrt{\lambda_k})$ and $M(-\sqrt{\lambda_k})$ and this is impossible by the definition of F_k. The same reasoning establishes the desired result for the g's. For the h's the situation is even simpler: of the three numbers $\lambda_n, \lambda_m, \lambda_k$ the one with highest subscript is included in the group generated by the other two, and therefore by the way V was defined for the h's the corresponding square root choices are compatible.

In order to be able to state our final theorem we need to define a concept introduced by A. Haar.[10] An orthogonal set $f_1, f_2, \cdots$ of bounded functions is said to be relatively complete if the function $f(x) \equiv 1$, and the functions f_nf_m, $n, m = 1, 2, \cdots$ belong to the closed linear manifold spanned by the f_j. In terms of this concept we can state our final theorem.

---

[10] A. Haar, "Über die Multiplikationstabelle der orthoganalen Funktionensysteme," *Mathematische Zeitschrift*, vol. 31 (1930), pp. 769-798.

THEOREM 6. *If T has pure point spectrum and if there exists a nowhere vanishing real valued characteristic function $n(x)$ belonging to the characteristic value -1, then a necessary and sufficient condition that T have a square root is that there exist a function $p(x)$ invariant under T and taking only the values ± 1, and a relatively complete orthogonal set $\phi_1, \phi_2, \cdots$ of characteristic functions such that the totality of the functions*

$$\{\phi_n(x), p(x)\phi_n(x)\,;\ n = 1, 2, \cdots\}$$

is a complete orthogonal set in L_2.

Remark. If we write $A = {}_x\{p(x) = +1\}$, $B = X - A$, then the condition says, essentially, that the transformation T behaves about the same way on the two invariant sets A and B. For, let P be the projection in L_2 on the manifold of all functions which vanish on A and let $Q = I - P$; write $\alpha_n = P\phi_n$, $\beta_n = Q\phi_n$. Then $\{\alpha_n\}$ and $\{\beta_n\}$ are complete orthogonal sets of characteristic functions of T on the manifolds $L_2(A)$ and $L_2(B)$ respectively. The fact that $\phi_n(x)$ is orthogonal to $p(x)\phi_n(x)$ implies that the linear operator W with domain $L_2(A)$ and range $L_2(B)$ defined by $W\alpha_n = \beta_n$ is isometric. Moreover the relative completeness of the set ϕ_n implies $W(\alpha_n\alpha_m) = \beta_n\beta_m$. By an easy modification of Lemma 1 this is seen to imply that there exists a one to one measure preserving transformation R of A on B such that for every $f \in L_2(A)$, $Wf(x) = f(Rx)$. Since the transformations T_A and T_B induced by T on A and B, respectively, have α_n and β_n as characteristic functions corresponding to the same characteristic value, it is easy to verify that $RT_AR^{-1} = T_B$.

Proof of necessity. Suppose that T has a square root, $T = S^2$. Write

$$C_0 = {}_x\{n(x) > 0\}, \quad C_1 = {}_x\{n(x) < 0\}.$$

Then $TC_i = C_{i+1}$, (where the subscripts are interpreted modulo 2). If we write $C_{ij} = C_i \cdot SC_j$, and $A = C_{00} + C_{11}$, $B = C_{01} + C_{10}$, then it follows, just as in the proof of Theorem 4, that $TA = SB = A$, $TB = SA = B$.

Let λ_n be any one of the characteristic values of U; let $\{a_i^n(x)\}$ be a complete orthogonal set of bounded functions in the manifold of (A, λ_n)-functions. Write $b_i^n(x) = a_i^n(Sx)/\lambda_n$: then $\{b_i^n(x)\}$ is a complete orthogonal set in the manifold of (B, λ_n)-functions, and together the $a_i^n(x)$ and $b_i^n(x)$ are a complete orthogonal set in the manifold of all characteristic functions of characteristic value λ_n. We write

$$f_i^n(x) = a_i^n(x) + b_i^n(x), \quad g_i^n(x) = p(x)f_i^n(x) = a_i^n(x) - b_i^n(x).$$

It is immediately verified that these functions are a complete orthogonal set in the manifold of characteristic functions of λ_n.

We shall now prove that the set of functions $f_i^n(x)$, $i, n = 1, 2, \cdots$ is a relatively complete set. Let $a(x)$ and $b(x)$ be the characteristic functions of A and B respectively; suppose that $\lambda_1 = 1$. Then since $a(x)$ is an (A, λ_1)-function we have $a(x) = \Sigma_a c_a a_a^1(x)$ and

$$b(x) = a(Sx) = \Sigma_a c_a a_a^1(Sx) = \Sigma_a c_a b_a^1(x),$$

so that

$$1 \equiv a(x) + b(x) = \Sigma_a c_a (a_a^1(x) + b_a^1(x)) = \Sigma_a c_a f_a^1(x).$$

Hence the function $f(x) \equiv 1$ is in the manifold of functions spanned by the f_i^n. If $f_i^n(x) f_j^m(x) \equiv 0$ then clearly this product is also in this manifold; if $f_i^n(x) f_j^m(x) \not\equiv 0$ then this product is a characteristic function belonging to the characteristic value λ_k say. We have

$$f_i^n f_j^m = (a_i^n + b_i^n)(a_j^m + b_j^m) = a_i^n a_j^m + b_i^n b_j^m.$$

Since, clearly, $a_i^n a_j^m$ is an (A, λ_k)-function, we have

$$a_i^n(x) a_j^m(x) = \Sigma_a c_a a_a^k(x)$$

so that

$$b_i^n(x) b_j^m(x) = \frac{a_i^n(Sx)}{\lambda_n} \cdot \frac{a_j^m(Sx)}{\lambda_m} = \Sigma_a c_a \frac{a_a^k(Sx)}{\lambda_k} = \Sigma_a c_a b_a^k(x).$$

It follows that

$$f_i^n f_j^m = \Sigma_a c_a (a_a^k + b_a^k) = \Sigma_a f_a^k.$$

This completes the proof of the necessity of our conditions.

Proof of sufficiency. Let $\lambda_1, \lambda_2, \cdots$ be all characteristic values of T and let a complete orthogonal set of bounded characteristic functions belonging to λ_n be $\{f_i^n(x), g_i^n(x)\}$, where $g_i^n(x) = p(x) f_i^n(x)$ and where the $f_i^n(x)$, $i, n = 1, 2, \cdots$ are a relatively complete set. (We may take, for n fixed, $f_i^n(x)$, $i = 1, 2, \cdots$, to be those characteristic functions of the set $\phi_1, \phi_2, \cdots$ which correspond to λ_n). Take arbitrary fixed values of i, j, n, m; if $\lambda_n \lambda_m$ is a characteristic value write $\lambda_k = \lambda_n \lambda_m$, otherwise write $\lambda_k = \lambda_1 = 1$. The hypotheses made imply the following relations:

$$f_i^n f_j^m = g_i^n g_j^m = \Sigma_a c_a f_a^k(x),$$

$$f_i^n g_j^m = g_i^n f_j^m = \Sigma_a c_a g_a^k(x).$$

We define an operator V by the equations:

$$V f_i^n = \left(\frac{1 + \lambda_n}{2}\right) f_i^n + \left(\frac{-1 + \lambda_n}{2}\right) g_i^n,$$

$$V g_i^n = \left(\frac{1 - \lambda_n}{2}\right) f_i^n + \left(\frac{-1 - \lambda_n}{2}\right) g_i^n.$$

It is easy to verify that V is a unitary operator whose square is U. We prove that $V(\phi\psi) = (V\phi)(V\psi)$ whenever ϕ and ψ are chosen from the set f_i^n, g_j^m. The proof is merely a matter of calculation; we give the steps for only one of three possible cases.

$$\begin{aligned}(Vf_i^n)(Vg_j^m) &= \left[\left(\frac{1+\lambda_n}{2}\right) f_i^n + \left(\frac{-1+\lambda_n}{2}\right) g_i^n\right] \\ &\qquad\cdot \left[\left(\frac{1-\lambda_m}{2}\right) f_j^m + \left(\frac{-1-\lambda_m}{2}\right) g_j^m\right] \\ &= (\Sigma_a c_a f_a^k)\left(\frac{1-\lambda_k}{2}\right) + (\Sigma_a c_a g_a^k)\left(\frac{-1-\lambda_k}{2}\right) \\ &= \Sigma_a c_a \left[\left(\frac{1-\lambda_k}{2}\right) f_a^k + \left(\frac{-1-\lambda_k}{2}\right) g_a^k\right] = V(f_i^n g_j^m).\end{aligned}$$

The theorem now follows from Lemma 1 and the remark (A) following it.

3. Examples. In this section we give some examples illustrating the scope of our theorems and demonstrating the necessity of the detailed proofs of Theorems 5 and 6.

Let X be a finite interval, $a \leqq x < b$, and γ an arbitrary real number. Denote by $T(a, b; \gamma)$ the transformation defined by $Tx = x + \gamma$, mod (a, b): i. e., the transformation that sends each point x, $a \leqq x < b$, into $y = x + \gamma + n(b - a)$, where n is the unique integer for which $a \leqq y < b$. $T(a, b; \gamma)$ is a one to one measure preserving transformation of (a, b) into itself; it is known that $T(a, b; \gamma)$ is metrically transitive if and only if $\gamma/(b - a)$ is irrational and that T has pure point spectrum with characteristic values $e^{2\pi i n\gamma/(b-a)}$ and corresponding characteristic functions $e^{2\pi i n(x-a)/(b-a)}$ for $n = 0, \pm 1, \pm 2, \cdots$.

(1) Let X be a finite interval, $a \leqq x < b$, and let $\gamma/(b - a)$ be an irrational number. Denote by $R = R(a, b; \gamma)$ the transformation defined as follows:

For $a \leqq x < (a + b)/2$, $Rx = T(a, b; (a + b)/2)x = x + (a + b)/2$;

for $(a + b)/2 \leqq x < b$, $Rx = T(a, (a + b)/2; \gamma)T(a, b; (a + b)/2)x$.

It is readily verified that R is a metrically transitive transformation with pure point spectrum (which is therefore simple) and that -1 is a characteristic value of R. Hence by Theorem 4, R does not have a square root.

(2) The most important property of metrically transitive transformations used in the proof of Theorem 3 is the fact that the characteristic values form a group, and in fact it is easy to extend the proof to all transformations with this property. The following is an example of a transformation which does not have this property. Let X be the interval $0 \leqq x < 1$ and γ an irrational number. Define S by the conditions

$$Sx = T(0, 1/2; \gamma)x \quad \text{for} \quad 0 \leqq x < 1/2,$$
$$Sx = T(1/2, 1; \gamma + (1/4))x \quad \text{for} \quad 1/2 \leqq x < 1.$$

It is easy to see that -1 is not in the spectrum of S, but $\pm e^{4\pi i\gamma}$ are both characteristic values.

(3) The above example had the property that X was decomposable into invariant sets on which the characteristic values form a group. To show that this is not the general case we consider the following example. Denote by $S(\gamma)$ the transformation in (2). Let X be the (countably) infinite dimensional product space of the unit interval with itself, with the usual multiplicative measure, and let $\gamma_1, \gamma_2, \cdots$ be a sequence of linearly independent real numbers. Define the transformation S' for a point $x \in X$, $x = (x_1, x_2, \cdots)$ by $S'x = (S(\gamma_1)x_1, S(\gamma_2)x_2, \cdots)$. In this case it is impossible to split X into a countable number of parts on each of which the characteristic values of S' form a group.

(4) An example which typifies the situation described in Theorem 6 is the following. Let X be the unit interval, and γ an irrational number. Let R' be defined by the relations

$$R'x = R(0, 1/2; \gamma)x \quad \text{for} \quad 0 \leqq x < 1/2,$$
$$R'x = R(1/2, 1; \gamma)x \quad \text{for} \quad 1/2 \leqq x < 1,$$

where $R(a, b; \gamma)$ is the transformation defined in (1). Let $p(x) = -1$ for $0 \leqq x < 1/2$, $p(x) = 1$ for $1/2 \leqq x < 1$. Let

$$\phi_n(x) = e^{8\pi i n x} \quad \text{for} \quad 0 \leqq x < 1/4 \quad \text{and} \quad 1/2 \leqq x < 3/4,$$
$$\phi_n(x) = e^{8\pi i n x}/e^{4\pi i n\gamma} \quad \text{for} \quad 1/4 \leqq x < 1/2 \quad \text{and} \quad 3/4 \leqq x < 1,$$
$$(n = 0, \pm 1, \pm 2, \cdots).$$

The functions $p(x)$ and $\phi_1, \phi_2, \cdots$ satisfy the hypotheses of Theorem 6, and therefore R' has a square root. One such square root Q that our method yields is defined as follows:

$$\begin{array}{lll} \text{For} & 0 \leqq x < 1/4, & Qx = T(0, 1; 1/2)x = x + 1/2, \\ \text{for} & 1/4 \leqq x < 1/2, & Qx = T(0, 1; 1/2)x = x + 1/2, \\ \text{for} & 1/2 \leqq x < 3/4, & Qx = T(0, 1; 3/4)x, \\ \text{for} & 3/4 \leqq x < 1, & Qx = T(0, 1/4; \gamma)T(0, 1; 1/4)x. \end{array}$$

4. Concluding remarks. (A) We observe that no sort of uniqueness could be hoped for in any general theorem that asserts the existence of a square root: witness, for instance, the identity transformation, which has any transformation of order two for a square root. The method of Theorem 5 (or of Theorem 3) yields a unique square root only if every characteristic value is the square of another characteristic value; in all other cases there are several square roots obtainable by that method. The unique square root which the method associates with the identity transformation is the identity itself; more generally if the characteristic values of a transformation form a finite group not containing -1 then the method of Theorem 5 yields as a square root a transformation with the same group of characteristic values. (This fact is most easily seen by an examination of the proof of Lemma 3.) Thus, for example, the square root, given by Theorem 5, of $T(0,1;1/3)$ is not $T(0,1;1/6)$ but $T(0,1;2/3)$.

(B) The square root of a transformation is more likely to be metrically transitive and less likely to be continuous than the transformation itself. For if $S^2 = T$, then any set invariant under S is invariant under T, but not conversely. If T is the transformation defined by $Tx = -x$ for $-1 \leqq x \leqq 1$ then a square root for T is easily found by Theorem 6, but since T and the identity are the only two continuous measure preserving transformations of $(-1,1)$ into itself, T has no continuous square roots.

(C) We remark, finally, on the rôle of the topological assumptions we made on X. They were necessary only in order to apply Lemma 1: a result which most likely remains true if all topological assumptions are removed, but which has never been proved without them. We can, however, get along without the topology even so, if we are willing to restrict our attention to set transformations. The sort of set correspondences we consider are, of course, one to one and measure preserving, and preserve (countable) sum, product, and complement. To any such transformation T there corresponds a unitary operator U determined by the condition that if F is the characteristic function of a measurable set E, then Uf is the characteristic function of TE. All our results may be applied to these unitary operators and our theorems yield, whenever possible, a unitary operator V satisfying (*) and $V^2 = U$. The advantage of considering set transformations (and the rôle of Lemma 1 and the topological assumptions necessary for its proof) becomes apparent only now: it takes Lemma 1 to assure us of the existence of a point transformation which generates V, but a set transformation which generates V always exists. (See reference in footnote (¹)).

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, NEW JERSEY.

POST ALGEBRAS. I. POSTULATES AND GENERAL THEORY.*

By PAUL C. ROSENBLOOM.

Historical introduction and summary of results. In 1920, Post[1] (1) gave a definition of n-valued logic that was a generalization of the usual 2-valued calculus. He defined the most important operations and discussed some of their properties by means of tables of values. Webb (2) reduced the required number of undefined ideas to one. Later Webb (3) proved most of the important propositions of the n-valued logic by the numerical interpretation in terms of congruences. This consists, essentially, in using tables of values in a general way that eliminates trial and error. The Lukasiewicz-Tarski (4) n-valued logics are included in the Post calculi, but not conversely. In fact, the Post logics are symbolically complete,[2] i. e., all operations and relations in these systems definable by tables of values are definable in terms of the primitive ideas of these logics; the Lukasiewicz-Tarski systems do not share this property. We shall confine our attention, in this paper, to the Post algebras,[3] not concerning ourselves with their interpretation as logics, but with the abstract algebraic properties of the systems.

We give here the first[4] postulate-set for Post algebras and prove the fundamental theorems from the assumptions. The Post algebras are here generalized in analogy to the extension from 2-element Boolean algebras to n-element Boolean algebras. We prove that two n-valued Post algebras with the same finite number of elements are simply isomorphic, and deduce, as a corollary, that our postulate-set is complete when a postulate as to the number of elements is added. The analogues to the prime elements in Boolean algebras are defined and a unique factorization theorem is proven. This leads to the theorem that an n-valued Post algebra with $m(\geq 1)$ prime elements has n^m elements, and also to a numerical interpretation analogous to the Sheffer

---

* Received November 6, 1940; Revised June 28, 1941.

[1] The numbers in parentheses refer to the bibliography at the end of the paper.

[2] This term is attributed by Webb to J. C. C. McKinsey.

[3] We believe this term most appropriate for the algebras described here, just as Boolean algebra seems the best term for the algebra corresponding to the two valued logic.

[4] Since this was written Rosser's paper (8) appeared. His "Axiom group I" corresponds to the present development. However, he has a propositional calculus whereas we have here an algebra. The former explicitly states its rules of deduction, while the latter is "embedded" in ordinary logic, which is assumed implicitly.

"Boolean numbers." It turns out that Boolean algebra is such a degenerate special case that most Boolean theorems generalize into several theorems and most Boolean operations into several Post operations. In fact, the first Post algebra which exhibits the typical properties of them all is the six-valued case.[5]

Post algebras should have many applications since every operation or relation defined on a finite range can be defined in terms of Post operations,[6] and many cases of multiply-isomorphic objects[7] can be dealt with simultaneously. It follows from results of Bernstein (5) that this is not true of congruences, i. e., we cannot in general get a definition of an object isomorphic to a given finite operation or relation in terms of arithmetical addition and multiplication with respect to some modulus.[8] Furthermore, algebraic manipulations in Post algebras are in many ways simpler than the corresponding arithmetical operations.

The primitive frame.[9] We take as undefined a class K of elements,[10] a binary[11] operation $+$, and a unary[12] operation $'$ and make the following assumptions about them:[13]

Undefined: K, $+$, $'$.

We make the following definitions in order to be able to state our postulates in convenient form. We take $n \geq 2$ to be some fixed positive integer.

D1. $p + q + r = (p + q) + r$. (Similarly for $p + q + r + s$, etc.)

D2. $$\sum_{i=a}^{b} f(i) = f(a) + \cdots + f(b).$$

D3. $$p^0 = p. \qquad p^{m+1} = (p^m)'.$$

---

[5] For example, P9 does not begin to show its general form until $n \geq 6$. Of course, this is true only with respect to the way in which the theorems are derived from the postulates and not for the class of theorems as such.

[6] See Post (1), Webb (3), or this paper T57, which implies as a corollary that every n-ary operation on a range of n elements can be defined in terms of the operations in an n-valued Post algebra.

[7] The definition alluded to in [5] will yield an operation multiply isomorphic to the given one defined on a range of n^m elements.

[8] If the range has a prime number of elements, it is possible to define an isomorphic operation or relation in terms of addition and multiplication with respect to a modulus; see Bernstein (5).

[9] This term, introduced by Curry, means the totality of undefined ideas and unproved propositions.

[10] In the logical interpretation these are the "propositions."

[11] Our $+$ is Post's and Webb's V_n.

[12] Our $(')$ is Post's and Webb's $\sim_n$.

[13] We use, for brevity, such symbols as $\rightarrow$ for "implies," ϵ, and $\exists$ without further comment.

D4. $1(p) = \sum_{m=0}^{n-1} p^m.$

D5. $2(p) = (1(p))'. \quad m(p) = (1(p))^{m-1}. \qquad (n > m \geq 2).$

D6. $0(p) = ((n-1)(p))' = (1(p))^{n-1}.$

D7. $p_1 = (\sum_{m=1}^{n-1} p^m)^{n-1}.$

D8. $p_m = ((p + 2(p))^{n-1} + p^m)^{n-1}. \qquad (2 \leq m \leq n-1).$

D9. $p_j{}^i = (p^i)_j.$

D10. $p_{-i} = p_1{}^{n-i+1}.$

$1(p), 2(p), \cdots$, etc., will be proved independent of p. (T2, T3). They correspond to the so-called "truth values" which appear in the interpretation as logic. $p_1, p_2, \cdots$ are what might be called "step-functions" in view of T33 and T34 below.

D11. $-p = \sum_{m=1}^{n-1} p_m{}^m.$

p' and $-p$ are two generalizations of the "complement of p" in Boolean algebra.[14] In the two-valued case they degenerate to that operation. In general however, they are distinct.

D12. $pq = p \cdot q = -((-p) + (-q)).$

D13. $pqr = (pq)r.$

P1. $p \in K \,.\rightarrow.\, p' \in K.$

P2. $p, q \in K \,.\rightarrow.\, p + q \in K.$

For convenience, we use $H(p, q, r, \cdots)$ to mean "$p, q, r, \cdots \in K$, and all their combinations mentioned in the conclusion are in K."[15]

P3. $H(p, q) \,.\rightarrow.\, p + q = q + p.$

P4. $H(p, q, r) \,.\rightarrow.\, p + (q + r) = (p + q) + r.$

P5. $H(p) \,.\rightarrow.\, p + p = p.$

P6. $1(p) = (1(p))^n = (0(p))'.$

P7. $H(p, q, r) \,.\rightarrow.\, p + qr = (p + q)(p + r).$

P8. $H(p, q) \,.\rightarrow.\, \sum_{m=0}^{n-1} pq^m = p.$

P8 is a generalization of the Boolean law $pq + pq' = p$.

---

[14] Our $-$ is Post's $\simeq_n$.

[15] The second part of the definition of H was inserted to facilitate the finding of possible independence examples.

P9. $$H(p) \,.\!\rightarrow\!.\, p = p_1 + 2(p)p_1^{n-1} + \sum_{i=3}^{n-3}\{i(p) + [i+1](p)\}p_{-i} + \{[n-2](p)\}p_1^{3} + \{[n-1](p)\}p_1^{2} + (0(p))p_1^{1}.$$

P9 is a generalization of the Boolean law $p = p + 0p'$. It will be simplified in T30.[16]

P10. $$H(p, A_0, \cdots, A_{n-1}) \,.\!\rightarrow\!.\, \Big(\sum_{i=0}^{n-1} A_i p_1^{\,i}\Big)' = \sum_{i=0}^{n-1} A'_i p_1^{\,i}.$$[17]

P10 is a generalization of the Boolean law $(A_0p + A_1p')' = A'_0p + A'_1p'$.

P11. $$(\exists p, q) \cdot p, q \in K \cdot p \neq q.$$

A system[18] which obeys the above postulates will be called a Post Algebra of type n, and will be designated by P_n. If K contains exactly k elements (by P11, $k \geq 2$), P_n will be said to have the order k and will be designated by $P_n(k)$.

We now develop the elementary, but fundamental, formal properties. (Henceforth, such hypotheses as $H(p, q)$ will be understood implicitly.)

T1. $$pq = qp. \qquad \text{(D12, P3).}$$

T2. $$1(p) = 1(q). \qquad \text{(D4, P8, T1, P3, P4, P8, D4).}$$

Proof. $$1(p) = \sum_{i=0}^{n-1} p^i = \sum_{i=0}^{n-1} \Big(\sum_{j=0}^{n-1} p^i q^j\Big) = \sum_{i=0}^{n-1} \Big(\sum_{j=0}^{n-1} q^j p^i\Big)$$
$$= \sum_{j=0}^{n-1} \Big(\sum_{i=0}^{n-1} q^j p^i\Big) = \sum_{j=0}^{n-1} q^j = 1(q).$$

T3. $$m(p) = m(q). \qquad (n-1 \geq m \geq 0) \qquad \text{(T2, D5, D6).}$$

D14. $$m = m(p). \qquad (n-1 \geq m \geq 0).$$

D14 is justified since, by T3, $m(p)$ is a unique element of P_n independent of p. From now on we shall use the notation of D14. Using this notation, the following theorems are mere restatements of D4, D5, D6, D8, P6, P9.

T4. $$1 = \sum_{i=0}^{n-1} p^i.$$

T5. $$m = 1^{m-1}. \qquad (n-1 \geq m \geq 1).$$

---

[16] This postulate combines T27 and T30 in view of T14; it is not necessary to work T31 into this postulate. These considerations explain the complicated structure of this postulate.

[17] For convenience we make the convention that with small Latin letters, subscripts are used in accordance with D7 and D8; with capital letters subscripts have the usual meaning. Thus here, $A_0, A_1, \cdots$ are any n elements of K, not the elements derived from a single element A in accordance with D7 and D8.

[18] That is, an ordered triple $(K, +, ')$.

T6. $$0 = 1^{n-1}.$$

T7. $$p_m = ((p+2)^{n-1} + p^m)^{n-1}. \qquad (2 \leq m \leq n-1).$$

T8. $$1^n = 1 = 0'.$$

T9. $$p = p_1 + 2p_1^{n-1} + \sum_{i=3}^{n-3} (i + [i+1])p_{-i} + [n-2]p_1^3 + [n-1]p_1^2 + 0p_1^1.$$

T10. $$p + 1 = 1 + p = 1.$$

Proof. $$1 + p = p + 1 = p + p + \sum_{i=1}^{n-1} p^i = p + \sum_{i=1}^{n-1} p^i = 1.$$

T10 is the analogue of the well-known Boolean law.

T11. $$(p^m)^k = p^{m+k}. \qquad \text{(Induction)}.[19]$$

T12. $$m^n = m. \qquad (\text{T5, T11, T8, T5}).$$

T13. *$m^k = [m+k]$, where $[m+k]$ is the constant element of K represented by the least non-negative residue of $m+k$ modulo n.*[20]

We shall continue this convenient use of $[m]$.

T14. $$m_1 = 0. \qquad (m \neq 1;\ 0 \leq m \leq n-1) \qquad (\text{D7, T13, T10, T6}).$$
$$m_{-k} = 0. \qquad (m \not\equiv k \ (\text{mod}\, n);\ 0 \leq m \leq n-1, 1 \leq k \leq n).$$

Proof. Let $m \geq 2$. $$m_1 = (\sum_{i=1}^{n-1} m^i)^{n-1} = (m^{n-m+1} + \Sigma' m^i)^{n-1}$$
$$= (1 + \Sigma' m^i)^{n-1} = 1^{n-1} = 0.$$

(Σ' indicates summation over all terms except m^{n-m+1}.)

If $m = 1$, $n - m + 1 = n$; therefore this reasoning does not apply.

If $m = 0$, $0' = 1$ occurs in the summation and the above reasoning is valid. $m_{-k} = m_1^{n-k+1} = [m + n - k + 1]_1 = [m - k + 1]_1$.

T15. $$p = p + pq = pq + p. \qquad (\text{P8, P5, P8}).$$

Proof. $$p = \sum_{i=0}^{n-1} pq^i = pq + \sum_{i=1}^{n-1} pq^i = pq + pq + \sum_{i=1}^{n-1} pq^i$$
$$= pq + p = p + pq.$$

T16. $$p = (p+q)p = p(p+q). \qquad (\text{T15, P7, P5}).$$

Proof. $$p = p + pq = (p+p)(p+q) = p(p+q).$$

T17. $$p = pp. \qquad (\text{T16, P5}).$$

T18. $$p = -(-p). \qquad (\text{T17, D12, P5}).$$

---

[19] T11 is just a notational theorem, i. e., it follows from D3 alone.

[20] Within square brackets, all operations are in ordinary arithmetic; otherwise they are always Post operations.

Proof. $p = pp = -((-p) + (-p)) = -(-p)$.

T19. $$p = p1 = 1p.$$ (T16, T10).

This theorem shows that, as in Boolean algebra, 1 is the unit element.

T20. $$p(qr) = (pq)r = pqr.$$ (D12, T18, P4).

Proof. $$\begin{aligned} p(qr) &= -\{-p + -(qr)\} = -\{-p + -(-\{-q + -r\})\} \\ &= -\{-p + (-q + -r)\} = -\{(-p + -q) + -r\} \\ &= -\{-(pq) + -r\} = (pq)r. \end{aligned}$$

T21. $$pq + pr = p(q + r).$$ (P7, T15, P7, T16).

Proof. $$\begin{aligned} pq + pr &= (pq + p)(pq + r) = p(pq + r) \\ &= p(p + r)(q + r) = p(q + r). \end{aligned}$$

T22. $-0 = 1$. (T9, T14, T17, T15, D11, D8, P5, T13, T10, T6, T10).

Proof. $$\begin{aligned} 1 &= 1_1 + 2 \cdot 0_1 + \sum_{i=3}^{n-3} (i + [i+1])1_i + [n-2]4_1 + [n-1]3_1 + 0 \cdot 2 \\ &= 1_1 + 2 \cdot 0 + \sum_{i=3}^{n-3} (i + [i+1])0 + [n-2]0 + [n-1]0 + 0 \cdot 0 \\ &= 1_1 + 2 \cdot 0 + \sum_{i=3}^{n-3} (i + [i+1])0 + [n-2]0 + [n-1]0 + 0 \\ &= 1_1 + 0. \end{aligned}$$

Therefore $$\begin{aligned} -0 &= \sum_{i=1}^{n-1} 0_i{}^i = 1_1 + 2_2 + \sum_{i=3}^{n-1} 0_i{}^i = 1_1 + ((2+2)^{n-1} + 2^2)^{n-1} + \sum_{i=3}^{n-1} 0_i{}^i \\ &= 1_1 + (2^{n-1} + 2^2)^{n-1} + \sum_{i=3}^{n-1} 0_i{}^i = 1_1 + (1 + 2^2)^{n-1} + \sum_{i=3}^{n-1} 0_i{}^i \\ &= 1_1 + 1^{n-1} + \sum_{i=3}^{n-1} 0_i{}^i = 1_1 + 0 + \sum_{i=3}^{n-1} 0_i{}^i = 1 + \sum_{i=3}^{n-1} 0_i{}^i = 1. \end{aligned}$$

I have been unable to find a simpler proof of this theorem.

T23. $$p0 = 0p = 0.$$ (D12, T22, T10, T18).

Proof. $p0 = -(-p + -0) = -(-p + 1) = -1 = -(-0) = 0$.

T24. $$p + q = -((-p)(-q)).$$ (D12, T18).

T25. $$p + 0 = 0 + p = p.$$ (T15, T23).

Proof. $p + 0 = p + 0p = p$.

As in Boolean algebra, 0 is the zero element.

T26. $$1_1 = 1.$$ (See proof of T22).

T27. $$i + [i+1] = i. \qquad (3 \leq i \leq n-3).$$ (T14, T19).

Proof. $i = i_1 + 2[n+i-1]_1 + \sum_{j=3}^{n-3} (j + [j+1]) i_{-j} + [n-2][i+3]_1$
$+ [n-1][i+2]_1 + 0[i+1]_1.$
$i_1 = 0 = [n+i-1]_1 = [i+3]_1 = [i+2]_1.$
$0 = i_{-j}$ except when $n+i-j+1 \equiv 1, \bmod n$, i. e., $i = j$.
Then, $i_{-j} = 1.$
$i = (i + [i+1]) \cdot 1 = i + [i+1].$

We are now able to simplify P9.

T28. $p = \sum_{i=2}^{n-1} i \cdot p_{-i} + 1 \cdot p_1{}^0 + 0 \cdot p_1{}^1.$ (T27).

T29. $p^n = p.$ (T28, P10, T12).

Proof. $p^n = \sum_{i=2}^{n-1} i^n p_{-i} + 1^n p_1{}^0 + 0^n p_1{}^1 = \sum_{i=2}^{n-1} i p_{-i} + 1 p_1{}^0 + 0 p_1{}^1 = p.$

T30. $p = \sum_{i=0}^{n-1} i p_1{}^{n-i+1} = \sum_{i=0}^{n-1} i p_{-i}.$

T31. $2 + 2' = 2.$ $(n \geqq 3).$ (T26, D7, T5, T29, P5).

Proof. $1 = 1_1 = (\sum_{i=1}^{n-1} 1^i)^{n-1}.$ $2' + 2 = 2' + 1' = 2' + (\sum_{i=1}^{n-1} 1^i)^n$
$= 2' + \sum_{i=1}^{n-1} 1^i = 2' + 2 + 2' + \sum_{i=4}^{n-1} i = 2 + 2' + \sum_{i=4}^{n-1} i$
$= \sum_{i=1}^{n-1} 1^i = 1' = 2.$

T32. $2 + i = 2.$ $(2 \leqq i \leqq n-1$, or $i = 0).$ Similar to above proof.

Now we are able to characterize the function p_k more precisely and show the agreement between our definition and Post's.

T33. $i_k = 0.$ $(i \neq 1;\ 0 \leqq i \leqq n-1, 1 \leqq k \leqq n-1).$
(T7, T32, T13, T10, T6).

Proof. $i_k = \{((i+2)^{n-1} + i^k\}^{n-1} = (2^{n-1} + i^k)^{n-1}$
$= (1 + i^k)^{n-1} = 1^{n-1} = 0.$

T34. $1_k = k.$ (T7, T10, T6, T25, T13).

Proof. $1_k = \{((1+2)^{n-1} + 1^k\}^{n-1} = (1^{n-1} + 1^k)^{n-1}$
$= (0 + 1^k)^{n-1} = (1^k)^{n-1} = 1^{k+n-1} = [k+n] = k.$

T35. $-i = [n-i+1].$ $(0 \leqq i \leqq n-1).$ (D11, T33, T34, T13).

Proof. $-i = \sum_{j=1}^{n-1} i_j{}^j.$ All the terms equal 0 except the one with $i^j = 1$, i. e., $j = [n-i+1]$. Hence, $-i = 1_j = j = [n-i+1].$

D15. $p \subset q . \leftrightarrow . p + q = q.$

T36. $p \subset q . \leftrightarrow . pq = p.$

Proof. $p + q = q . \rightarrow . p(p + q) = p = pq.$
$p = pq . \rightarrow . p + q = pq + q = q.$

T37–T44 are proved by D15 or T36.

T37. $p \subset q . \rightarrow . p + r \subset q + r$ *and* $pr \subset qr.$

T38. $p \subset p + q. \quad pq \subset p.$

T39. $p \subset p.$

T40. $p \subset q$ *and* $q \subset r . \rightarrow . p \subset r.$

T41. $p \subset q$ *and* $p \subset r . \rightarrow . p \subset qr.$

T42. $p \subset r$ *and* $q \subset r . \rightarrow . p + q \subset r.$

T43. $0 \subset p \subset 1.$

T44. $p \subset q$ *and* $q \subset p . \rightarrow . p = q.$

T45. $p \subset q . \leftrightarrow . -q \subset -p.$ (D15, T24, T36).

Proof. $p \subset q . \leftrightarrow . p + q = q . \leftrightarrow . (-p)(-q) = -q . \leftrightarrow . -q \subset -p.$

T46. $i \subset 2.$ $(2 \leq i \leq n - 1).$ (T32).

T47. $[n-1] = -2 \subset i.$ $(1 \leq i \leq n - 1).$ (T45, T46).

T48. $1 \leq i \leq j \leq [n-1] : \rightarrow : j \subset i. \quad i + j = i. \quad ij = j.$ (T39, D15, T27, T46, T47).

Proof. $1 \leq i = j \leq [n-1] . \rightarrow . j \subset i.$

If $1 \leq i \leq j \leq [n-1]$, and $j \subset i$, and $[j+1] \leq [n-1]$, then $1 \leq i < [j+1] \leq [n-1]$, and $[j+1] \subset j$. Therefore $[j+1] \subset i$. Therefore the theorem is true by induction on j.

T49. $1 \leq i \leq j \leq [n-1] . \rightarrow . j_1 \subset i_1.$ (T26, T14, T43).

Proof. $j = 1$ and hypothesis $. \rightarrow . j = i = 1 = j_1 = i_1 . \rightarrow . j_1 \subset i_1.$
$j \neq 1$ and hypothesis $. \rightarrow . j_1 = 0 . \rightarrow . j_1 \subset i_1.$

T50. $0 \leq i, j \leq [n-1] . \rightarrow . (i + j)_1 = i_1 + j_1.$

Proof. Either $i \leq j$ or $j \leq i$. Assume $i \leq j$. If $i = 0$, $(i + j)_1 = j_1 = i_1 + j_1$. If $i \neq 0, 1 \leq i \leq j \leq [n-1]$. Therefore $j \subset i$. $j + i = i$. $j_1 \subset i_1$. $j_1 + i_1 = i_1$. Therefore $(i + j)_1 = i_1 = i_1 + j_1$. By symmetry, the theorem is true for $j \leq i$.

The fundamental theorem in Post algebras. This is the analogue of the Boolean theorem $f(p) = f(1)p + f(0)p'$. Before we can prove the theorem in general, we must deduce certain special cases.

T51. $$p^i = \sum_{j=0}^{n-1} j^i p_{-j} = \sum_{j=0}^{n-1} [j+i] p_{-j}. \qquad \text{(P10, T13).}$$

T52. $$\sum_{j=0}^{n-1} p_1^{\,i} = 1. \qquad \text{(T4, T21, T4).}$$

Proof. $1 = \sum_{i=0}^{n-1} p^i = \sum_{i=0}^{n-1} \left(\sum_{j=0}^{n-1} j^i p_{-j} \right) = \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} j^i \right) p_{-j} = \sum_{j=0}^{n-1} p_{-j} = \sum_{j=0}^{n-1} p_1^{\,n-j+1}.$

T53.[21] $$f(p) = \sum_{i=0}^{n-1} f(i) p_{-i} = \sum_{i=0}^{n-1} f(-i) p_1^{\,i}.$$

Proof. For $f(p) = p$, it is true by T30. For $f(p) = a$, a constant, we have by T52, $a = a \cdot 1 = a \sum_{i=0}^{n-1} p_1^{\,i} = \sum_{i=0}^{n-1} a p_1^{\,i}$. If T53 is true for $f(p)$, then for $g(p) = (f(p))'$,

$$g(p) = (f(p))' = \left(\sum_{i=0}^{n-1} f(i) p_{-i} \right)' = \sum_{i=0}^{n-1} (f(i))' p_{-i} = \sum_{i=0}^{n-1} g(i) p_{-i}.$$

If T53 is true for $f(p)$ and $g(p)$, then for $h(p) = f(p) + g(p)$,

$$h(p) = \sum_{i=0}^{n-1} f(i) p_{-i} + \sum_{i=0}^{n-1} g(i) p_{-i} = \sum_{i=0}^{n-1} (f(i) + g(i)) p_{-i} = \sum_{i=0}^{n-1} h(i) p_{-i}.$$

By induction, then, T53 is true for all $f(p)$ expressible in terms of the undefined ideas.

T54. $$f(p) = \sum_{i=0}^{n-1} A_i p_{-i} \textit{ for all } p \,.\!\rightarrow.\, A_i = f(i).$$

Proof. $f(i) = \sum_{j=0}^{n-1} A_j i_{-j} = A_i$, by T13, T14, T26. Hence the expansion is unique.

We now establish some important consequences.

T55. $$f\left(\sum_{i=0}^{n-1} A_i p_{-i} \right) = \sum_{i=0}^{n-1} f(A_i) p_{-i}.$$

Proof. Let $g(q) = \sum_{i=0}^{n-1} A_i q_{-i}$; $h(q) = f(g(q))$, for all q. Then

$$h(p) = \sum_{i=0}^{n-1} h(i) p_{-i} = \sum_{i=0}^{n-1} f(g(i)) p_{-i} = \sum_{i=0}^{n-1} f(A_i) p_{-i}.$$

T56. $$\left(\sum_{i=0}^{n-1} A_i p_{-i} \right) \left(\sum_{i=0}^{n-1} B_i p_{-i} \right) = \sum_{i=0}^{n-1} A_i B_i p_{-i}.$$

Proof. Let $h(x) = \left(\sum_{i=0}^{n-1} A_i x_{-i} \right) \left(\sum_{i=0}^{n-1} B_i x_{-i} \right)$ for all x. Then

---

[21] Throughout this paper $f(p)$, $g(p)$, etc., stand for functions expressible in terms of the primitives, $+$ and $\cdot$.

$$h(p) = \sum_{i=0}^{n-1} h(i) p_{-i} = \sum_{i=0}^{n-1} A_i B_i p_{-i}.$$

T57. $$F(p, q, \cdots, s) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \cdots \sum_{k=0}^{n-1} F(i, j, \cdots, k) p_{-i} q_{-j} \cdots s_{-k}.$$

Proof. The theorem is true for functions of 1 variable. If the theorem is true for functions of N variables, and $F(p, q, \cdots, s)$ is a function of $N + 1$ variables, then

$$F(p, q, \cdots, s) = \sum_{i=0}^{n-1} F(i, q, \cdots, s) p_{-i}; \text{ but}$$
$$F(i, q, \cdots, s) = \sum_{j, \ldots, k}^{n-1} F(i, j, \cdots, k) q_{-j} \cdots s_{-k}.$$

Therefore

$$F(p, q, \cdots, s) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \cdots \sum_{k=0}^{n-1} F(i, j, \cdots, k) p_{-i} q_{-j} \cdots s_{-k}.$$

As before the coefficients are unique.

Theorems T51–T57 *hold equally well if* $+$ *and* $\cdot$ *are the arithmetical operations modulo* n *and* K *is the set of residue classes modulo* n, *provided that the* p_i*'s are added as primitives (even this will be unnecessary if* n *is a prime number.)* To see this, let $\overset{b}{\underset{i=a}{S}} F(i)$ be arithmetical summation.

Let $g(p) = \overset{n-1}{\underset{i=0}{S}} f(i) \times p_{-i}$ (here $a \times b$ is arithmetical multiplication).

Then $$\begin{aligned} g(m) &= \overset{n-1}{\underset{i=0}{S}} f(i) \times m_{-i} \\ &= f(m) \times m_{-m} && \text{(by T14).} \\ &= f(m) \times 1 = f(m). \end{aligned}$$

Therefore $$\begin{aligned} g(p) &= \sum_{i=0}^{n-1} g(i) p_{-i} && \text{(Post operations!—by T53).} \\ &= \sum_{i=0}^{n-1} f(i) p_{-i} = f(p) && \text{(by T53).} \end{aligned}$$

Therefore $f(p) = \overset{n-1}{\underset{i=0}{S}} f(i) \times p_{-i}.$

From this the analogues in arithmetical operations of T51–T57 follow almost immediately.[22]

T58. $$p_k{}^i = kp_1{}^j. \qquad \text{(T53, T33, T34).}$$

Proof. $$p_k{}^j = \sum_{i=0}^{n-1} i_k{}^j p_{-i} = \sum_{i=0}^{n-1} [i + j]_k p_{-i}$$
$$= kp_{-(n+1-j)} = kp_1{}^{n+1-(n+1-j)} = kp_1{}^j.$$

---

[22] This observation is due to the referee.

T59. $f(i)$ is $0, 1, 2, \cdots,$ *or* $[n-1]$ *for* $i = 0, 1, \cdots, [n-1]$.[23]

Proof. This follows easily by induction as in T53.

T60.
$$f(p) = \sum_{i=0}^{n-1} A_i p_k^{n-i+1} \textit{ for all } p : \rightarrow : f(i) \subset k.$$
$$f(i) = k, [k+1], \cdots, [n-1], \textit{ or } 0.$$

Proof. $$f(p) = \sum_{i=0}^{n-1} A_i p_k^{n-i+1} = \sum_{i=0}^{n-1} k A_i p_1^{n-i+1} \quad (\text{all } p).$$

Therefore (T54, T38) $\quad f(i) = kA_i \subset k.$

Therefore (T48) $\quad f(i) = k, [k+1], \cdots, [n-1],$ or 0.

Conversely, we have

T61.
$$f(i) \subset k \ (\textit{all } i) : \rightarrow : f(p) = \sum_{i=0}^{n-1} f(i) p_k^{n-i+1}.$$

Proof. $$f(p) = \sum_{i=0}^{n-1} f(i) p_{-i} = \sum_{i=0}^{n-1} f(i) p_1^{n-i+1}.$$

But $\quad f(i) = f(i)k. \qquad$ (T36).

Therefore $$f(p) = \sum_{i=0}^{n-1} f(i) k p_1^{n-i+1} = \sum_{i=0}^{n-1} f(i) p_k^{n-i+1}.$$

T62.
$$i \not\equiv j \bmod n \,.\rightarrow.\, p_1^i p_1^j = 0.$$

Proof. $m_1^i m_1^j = [m+i]_1 [m+j]_1 = 0$, by T14, since at most one of $[m+i]$ and $[m+j]$ can equal 1. The theorem follows by T53.

The question arises whether the functions p_1^i are determined by the properties in T52 and T62. The answer is given in T65 below. First we need two lemmas.

T63.
$$1 \neq 0.$$

Proof. $\quad 1 = 0 \,.\rightarrow.\, p = 1p = 0p = 0$, contradicting P11.

T64.
$$1^i = 1^j \,.\rightarrow.\, i \equiv j \bmod n.$$

Proof.
$$\begin{aligned} 1^i = 1^j \,.\rightarrow.\, 1 &= 1^{i+n-i} = 1^{j+n-i} \\ .\rightarrow.\, 1 &= 1_1 = 1_1^{j+n-i} = [j+n-i+1]_1 = 0 \\ &\text{if } j \not\equiv i \bmod n. \end{aligned}$$

T65. *If n distinct functions* $\phi_i(p)$, $i = 0, 1, 2, \cdots, [n-1]$, *none of which is identically* 0, *are built up from p by the primitive functions* $+$ *and* $'$ *and satisfy the relations*

---

[23] Here in T60, and in T65, $f(p)$ and $\phi_i(p)$ are built up by $+$ and $'$ from p only without the introduction of constant elements. Of course, they fail for functions like $f(p) = ap$, where a is not of the form 1^i.

$$(1)\ \sum_{i=0}^{n-1} \phi_i(p) = 1 \quad \textit{and} \quad (2)\ \phi_i(p)\phi_j(p) = 0 \quad (i \neq j) \ \textit{for all } p,$$

then the $\phi_i(p)$'s are the $p_1{}^{i}$'s in some order.[24]

Proof. Let $k = 1, 2, \cdots, [n-1]$, or 0. Then

$$\sum_{i=0}^{n-1} \phi_i(k) = 1. \qquad \phi_i(k) = 1, 2, \cdots, [n-1], \text{ or } 0 \ (\text{all } i).$$

If $\quad \phi_i(k) \neq 1$ for all i, $\quad \phi_i(k) \subset 2$ for all i. $\qquad$ (T46).

Therefore $1 = \sum_{i=0}^{n-1} \phi_i(k) \subset 2$ (T42). Therefore $1 = 2$, which is impossible. Therefore one of the $\phi_i(k)$'s, say $\phi_{i_k}(k)$, is equal to 1. If $j \neq i_k$,

$$\phi_j(k) = \phi_j(k) \cdot 1 = \phi_j(k)\phi_{i_k}(k) = 0, \text{ by } (2).$$

Therefore $\phi_i(k) = 0$ or 1 (all i, k). Since none of the ϕ's is identically 0, each is 1 for at least one value of k, and we have shown that not more than one can be 1 for a given value of k. Hence each is 1 for exactly one value of k and 0 for all others. That is $\phi_{i_k}(k) = 1$. $\phi_{i_k}(j) = 0 \ (j \neq k)$. Therefore $\phi_{i_k}(p) = \sum_{m=0}^{n-1} \phi_{i_k}(m) p_1{}^{n-m+1}$, whereupon the theorem follows.[25]

T66. $\qquad (p + q)_1 = p_1 + q_1.$

Proof.

$$(p+q)_1 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i+j)_1 p_{-i} q_{-j}$$
$$= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i_1 + j_1) p_{-i} q_{-j} = p_1 + q_1.$$

T67. $\qquad p \subset q \,.\!\rightarrow\!.\, p_1 \subset q_1.$

T68. $\qquad (p_1)_1 = p_1. \qquad$ (T55, T26, T14).

Proof.

$$(p_1)_1 = (1p_1 + 0p_1{}^{n-1} + 0p_1{}^{n-2} + \cdots)_1$$
$$= 1_1 p_1 + 0_1 p_1{}^{n-1} + 0_1 p_1{}^{n-2} + \cdots = p_1.$$

The Structure of Finite Post Algebras.

Prime elements and the unique factorization theorem.

D16. *p is prime ($p \in Pr$) $: \leftrightarrow : p \neq 1$ and $p \subset q : \rightarrow$ (all q) $: q = p$ or $q = 1$.*

---

[24] This fails without the restriction in [23]. For example, in a P_n with n prime elements (see D16), $Q_1, Q_2, \cdots Q_n$,[17] let $\phi_i(p)$ be the constant element $(Q_i)_1{}^{n-1}$. By T82, T90, T83 below, the hypotheses but not the conclusion of this theorem would be satisfied.

[25] The material in T58-T65 is due to the referee.

This definition is fundamental in the rest of the paper. In lattice-theoretic terminology, a prime is an element covered by 1. For $n = 2$, it reduces to Huntington's definition of "prime" or "irreducible" elements in a Boolean algebra.

From T69 on, we assume that K has a finite number of elements. All the previous developments are independent of this assumption.

T69. $$q \epsilon K \cdot q \neq 1 : \rightarrow : (\exists p) \cdot p \epsilon Pr \cdot q \subset p.$$

Proof. Hypothesis and $q \epsilon Pr : \rightarrow : q \epsilon Pr.$ $q \subset q.$ Hypothesis and

$$q \bar{\epsilon} Pr : \rightarrow : (\exists A_1) \cdot q \subset A_1 \cdot q \neq A_1 \cdot A_1 \neq 1.$$

If $A_1 \epsilon Pr$, then the theorem is proved. If $A_1 \bar{\epsilon} Pr$, then

$$(\exists A_2) \cdot A_1 \subset A_2 \cdot A_1 \neq A_2 \cdot A_2 \neq 1 \cdot A_2 = q \,.\rightarrow .\, q \subset A_1 \subset q.$$

Therefore $A_1 = q$. Therefore $A_2 \neq q$. Continuing we get a chain which must end since K is finite. Therefore the last A is a prime.

T70. $$p, q \epsilon Pr \text{ and } p \neq q \,.\rightarrow .\, p + q = 1 = p_1 + q_1.$$

Proof. $p \subset p + q.$ $q \subset p + q.$ Therefore $p + q = 1$ or p. If

$$p + q = p \text{ then } q \subset p, p \neq q, p \neq 1.$$

Therefore $q \bar{\epsilon} Pr$. Therefore $p + q \neq p$. $p + q = 1$. Therefore

$$p_1 + q_1 = (p + q)_1 = 1_1 = 1. \qquad \text{(T66).}$$

T71. $$p \epsilon Pr \,.\rightarrow .\, p = p_1 + 2 = p_1 + 2 \sum_{i=0}^{n-1} p_{-i}.$$

Proof. $$p = \sum_{i=0}^{n-1} i p_{-i} \subset p_1 + \sum_{\substack{i=0 \\ i \neq 1}}^{n-1} 2 p_{-i} \qquad \text{(T46).}$$

$$\subset p_1 + \sum_{i=0}^{n-1} 2 p_{-i} = p_1 + 2.$$

Therefore $p_1 + 2 = p$ or 1. (D16).

$$p_1 + 2 = 1 \,.\rightarrow .\, p_1 = (p_1)_1 + 2_1 = (p_1 + 2)_1 = 1_1 = 1. \qquad \text{(T66).}$$

$$.\rightarrow .\, p = p_1 + \sum_{i=2}^{n-1} i p_{-i} = 1 \qquad \text{(T10).}$$

$$.\rightarrow .\, p \bar{\epsilon} Pr.$$

Therefore $p_1 + 2 \neq 1$. Therefore $p_1 + 2 = p$.

T72. $$p \epsilon Pr : \rightarrow : p = p + 2. \qquad 2 \subset p.$$

Proof. Hypothesis $.\rightarrow .\, p + 2 = p_1 + 2 + 2 = p_1 + 2 = p.$

T73. $p \epsilon Pr : \rightarrow : p_1^{n-1+1} = p_{-i} = 0.$ $(i \neq 1, 2).$ $\rightarrow .\, p_1^{i} = 0.$ $(i \neq 0, n-1).$

Proof. Hypothesis $.\rightarrow. p = p_1 + 2\sum_{i=0}^{n-1} p_1^i$

$.\rightarrow. p^j = 1^j p_1 + 2^j \sum_{i=1}^{n-1} p_1^i$ (P10).

$.\rightarrow. p_1^j = 1_1^j p_1 + 2_1^j \sum_{i=1}^{n-1} p_1^i$ (T55).

$= [j+1]_1 p_1 + [j+2]_1 \sum_{i=1}^{n-1} p_1^i$

$= 0$ unless $[j+1] = 1$ or $[j+2] = 1$,

i. e., $j = n - 1$ or $j = 0$.

T74. $p \epsilon Pr :\rightarrow: p = p_1 + 2p_1^{n-1}$. $f(p) = f(1)p_1 + f(2)p_1^{n-1}$.

$p_1 + q = p_1 + qp_1^{n-1}$. $f(ap_1 + bp_1^{n-1}) = f(a)p_1 + f(b)p_1^{n-1}$.

T75. $p \epsilon Pr .\rightarrow. 1 = p_1 + p_1^{n-1}$.

Proof. Hypothesis $.\rightarrow. 1 = \sum_{i=0}^{n-1} p_1^i = p_1^0 + p_1^{n-1} = p_1 + p_1^{n-1}$.

We now consider an important subalgebra which is also a P_n (as we prove below).

D17. $q \epsilon \{p\} .\leftrightarrow. p_1 \subset q$.

D18. $q^* = p_1 + q'$.

D19. $q^{*1} = q^*$. $q^{*(k+1)} = (q^{*k})^*$.

We wish to prove that $\{p\}$ is a sub-P_n if p is a prime element, and to characterize it precisely.

T76. $p_1 + f(p_1 + q) = p_1 + f(q)$.

Proof. $p_1 + f(p_1 + q) = p_1 + f(p_1 + q\sum_{i=0}^{n-1} p_1^{n-i+1})$

(letting Σ' indicate summation with omission of $i = 1$)

$$= p_1 + f(1p_1 + qp_1 + q\Sigma' p_{-i}) = p_1 + f(1p_1 + q\Sigma' p_{-i})$$
$$= p_1 + f(1)p_1 + f(q)\Sigma' p_{-i} = p_1 + f(q)p_1 + f(q)\Sigma' p_{-i}$$
$$= p_1 + f(q)\sum_{i=0}^{n-1} p_{-i} = p_1 + f(q).$$

T77. $q^{*k} = p_1 + q^k$.

Proof. $q^{*1} = p_1 + q^1$.

If $q^{*k} = p_1 + q^k$,

$q^{*(k+1)} = (p_1 + q^k)^* = p_1 + (p_1 + q^k)' = p_1 + q^{k+1}$. (T68).

T78. $q \epsilon \{p\} .\rightarrow. q^{*n} = q$.

Proof. Hypothesis $.\rightarrow. q^{*n} = p_1 + q^n = p_1 + q = q$.

T79.
$$1 = \sum_{i=0}^{n-1} q^{*i}.$$

Proof. $\sum_{i=0}^{n-1} q^{*i} = \sum_{i=0}^{n-1} (p_1 + q^i) = p_1 + \sum_{i=0}^{n-1} q^i = p_1 + 1 = 1.$ (P5).

T80.
$$q \epsilon \{p\} \cdot p \epsilon Pr : \rightarrow : q = 1 \ or \ q \subset p.$$

Proof. Hypothesis $: \rightarrow : p_1 \subset q$

$$: \rightarrow : p = p_1 + 2 \subset q + 2$$
$$: \rightarrow : q + 2 = p \text{ or } q + 2 = 1.$$
$$q + 2 = p . \rightarrow . q \subset p = q + 2. \qquad \text{(T38).}$$

Therefore not

$$(q \subset p) : \rightarrow : q + 2 = 1 . \rightarrow . q_1 = q_1 + 2_1 = 1_1 = 1. \qquad \text{(T66).}$$
$$: \rightarrow : q = q_1 + \sum_{i=2}^{n-1} i q_{-i} = 1. \qquad \text{(T10).}$$

T81.
$$q \epsilon \{p\} \cdot p \epsilon Pr : \leftrightarrow : (\exists j) \cdot q = p_1 + 1^j.$$

Proof. $q = 1 : \rightarrow : q = p_1 + 1^0 \cdot q \epsilon \{p\}.$

$$q \epsilon \{p\} \cdot q^{*j} = 1 : \rightarrow : q = q^{*n}$$
$$= (q^{*j})^{*(n-j)} = 1^{*(n-j)} = p_1 + 1^{n-j}. \qquad \text{(T77).}$$
$$q \epsilon \{p\} \cdot q^{*j} \neq 1 \text{ for all } j : \rightarrow : \text{for all } j, q^{*j} \subset p.$$
$$: \rightarrow : 1 = \sum_{j=0}^{n-1} q^{*j} \subset p$$
$$: \rightarrow : p = 1.$$

Therefore $p \epsilon Pr \cdot q \epsilon \{p\} : \rightarrow : (\exists j) \cdot q^{*j} = 1$

$$: \rightarrow : (\exists j .) \cdot q = p_1 + 1^{n-j}.$$

This theorem is important in what follows, since it delimits $\{p\}$ so precisely.

T82. *If* $p, q, r, \cdots$ *are all the primes in* K, $p_1^{n-1} + q_1^{n-1} + r_1^{n-1} + \cdots = 1.$

Proof. Hypothesis and $a = p_1^{n-1} + q_1^{n-1} + \cdots \neq 1 : \rightarrow :$ there is a prime, say p, such that $a \subset p$. Therefore

$$a + p_1 \subset p + p_1 = p_1 + 2 + p_1 = p_1 + 2 = p.$$

But

$$a + p_1 = p_1 + p_1^{n-1} + q_1^{n-1} + \cdots = 1 + q_1^{n-1} + \cdots = 1.$$

Therefore $p = 1$. But $p \epsilon Pr$. Therefore $a = 1$.

T83. *Hypothesis* (T82) $. \rightarrow . 0 = p_1 q_1 \cdots .$

Proof. Hypothesis $. \rightarrow . p_1 q_1 \cdots = 1 p_1 q_1 \cdots$

$$\begin{aligned} &= (p_1^{n-1} + q_1^{n-1} + \cdots) p_1 q_1 \cdots \\ &= p_1^{n-1} p_1 q_1 \cdots + q_{-1}^{n-1} p_1 q_1 \cdots + \cdots \\ &= 0 + 0 + \cdots = 0. \qquad \text{(T62).} \end{aligned}$$

T84. $a \epsilon K$. *Hypothesis* (T82)

$$\begin{aligned} &: \rightarrow : a = (p_1 + a)(q_1 + a) \cdots \\ &: \rightarrow : (\exists i, j, \cdots) \cdot a = (p_1 + 1^i)(q_1 + 1^j) \cdots. \end{aligned}$$

Proof. Hypothesis $. \rightarrow . a = a + p_1 q_1 \cdots$

$$= (a + p_1)(a + q_2) \cdots. \qquad \text{(P7).}$$

$$a + p_1 \epsilon \{p\}, \qquad a + q_1 \epsilon \{q\}, \cdots.$$

Therefore $(\exists i, j, \cdots) \cdot p_1 + a = p_1 + 1^i$,

$$q_1 + a = q_1 + 1^j, \cdots \cdot a = (p_1 + 1^i)(q_1 + 1^j) \cdots. \qquad \text{(T81).}$$

We can put T84 in a more striking form if we define:

D20. *All elements of the form $p_1 + 1^i$, where p is a prime, are called powers of p*; (by T81, all powers of p are members of $\{p\}$, and conversely). T84 now states:

Every element a of K can be factored into powers of primes. We now prove that this factorization is essentially unique.

T85. $$(p_1 + q)^{*j} = p_1 + (p_1 + q)^j = p_1 + q^j. \qquad \text{(T77, T76).}$$

T86. *Hypothesis* (T82) *and if $p, q, r, \cdots$ are all distinct, and if $a = (p_1 + 1^i)(q_1 + 1^j) \cdots = (p_1 + 1^m)(q_1 + 1^t) \cdots$, then $1^i = 1^m$, $1^j = 1^t$, etc.*

Proof. $$p_1 + a = (p_1 + p_1 + 1^i)(p_1 + q_1 + 1^j) \cdots \qquad \text{(T70).}$$

$$= (p_1 + 1^i) 1 1 \cdots = p_1 + 1^i = p_1 + 1^m.$$

(By symmetry).

Therefore

$$\begin{aligned} (p_1 + a)^{*(n-i)} &= (p_1 + 1^i)^{*(n-i)} = (p_1 + 1^m)^{*(n-i)} \\ &= p_1 + (1^i)^{n-i} = p_1 + (1^m)^{n-i}. \end{aligned}$$

Therefore

$$1 = p_1 + 1 = p_1 + 1^n = p_1 + (1^i)^{n-i} = p_1 + (1^m)^{n-i} = p_1 + 1^{m+n-i}.$$

Therefore

$$1 = 1_1 = (p_1)_1 + 1_1^{m+n-i} = p_1 + 1_1^{m+n-i}. \qquad \text{(T66).}$$

$$m + n - i \not\equiv 0 \bmod n . \rightarrow . 1_1^{m+n-i} = 0 . \rightarrow . p_1 = 1 . \rightarrow . p = p_1 + 2 = 1.$$

Therefore

$$m + n - i \equiv 0 \bmod n. \qquad m \equiv i \bmod n. \qquad 1^m = 1^i.$$

Similarly $1^i = 1^j$, etc.

Therefore, every element a of K can be factored into powers of distinct primes in one and only one way, disregarding the order of the factors.

Further theorems and a numerical interpretation.

T87. *The elements* $0, 1, \cdots, (n-1)$ *are a* $P_n(n)$ *with respect to* $+$, $(')$.

Since they are elements of K, they obey P3–P10. But by T13, T48, and T25, and T63, P1, P2, and P11 are also satisfied. Obviously, by T64 this algebra is of Type $P_n(n)$.

T88. *If* $p \neq 1$, *then* $p_1 + 0, p_1 + 1, \cdots, p_1 + [n-1]$ *form a* $P_n(n)$ *with respect to* $+$, $(^*)$. *This is isomorphic to the one in* T87 *under the correspondence* $T(1^i) = p_1 + 1^i$.

Proof. By this correspondence, each element 1^i corresponds to a unique element $p_1 + 1^i$. Conversely, if

$$p_1 + 1^i = T(1^i) = T(1^j) = p_1 + 1^j,$$

then

$$\begin{aligned} 1 = p_1 + (1^i)^{n-i} &= p_1 + (p_1 + 1^i)^{n-i} \\ &= p_1 + (p_1 + 1^j)^{n-i} = p_1 + 1^{j+n-i}. \qquad \text{(T85)}. \end{aligned}$$

Therefore

$$1 = 1_1 = (p_1)_1 + 1_1{}^{j+n-i} = p_1 + 1_1{}^{j+n-i}.$$

$$j \not\equiv i \bmod n \,.\!\rightarrow\!.\, 1_1{}^{j+n-i} = 0 \,.\!\rightarrow\!.\, 1 = p_1 \subset p.$$

Hence $p = 1$. Therefore $j \equiv i \bmod n$. Therefore $1^i = 1^j$. Therefore, T generates a one-to-one correspondence.

$$T((1^i)') = T(1^{i+1}) = p_1 + 1^{i+1} = (p_1 + 1^i)^* = T(1^i)^*$$

$$1^i + 1^j = 1^i \text{ or } 1^j, \text{ say } 1^i.$$

Then

$$\begin{aligned} T(1^i + 1^j) = T(1^i) &= p_1 + 1^i \\ &= p_1 + 1^i + p_1 + 1^j = T(1^i) + T(1^j). \end{aligned}$$

Therefore T is an isomorphism. Therefore the set of elements $p_1 + 1^i$ is a $P_n(n)$ with respect to $+$, $(^*)$.

T89. *If p is a prime in K, then p, in that case equal to $T(2)$, is the only prime element of the subalgebra $\{p\}$.*

Proof. $q \in \{p\} \cdot p \in Pr$ (in K) $\cdot\, p \subset q :\rightarrow: q = 1$ or $q = p$. $:\rightarrow: p$ is a prime in $\{p\}$.[26] By T80, there can be no other.

In all the rest of this paper, $p, q, r, \cdots$ are all the distinct primes of K.

T90. $$p_1 = q_1^{n-1} + r_1^{n-1} + \cdots. \qquad p_1^{n-1} = q_1 r_1 \cdots.$$

Proof. $$p_1 + q_1 r_1 \cdots = (p_1 + q_1)(p_1 + r_1) \cdots. \qquad \text{(P7, T70)}.$$

Let $a = q_1 r_1 \cdots$. Therefore $p_1 + a = 1$. Therefore

$$p_1^{n-1} = p_1^{n-1} a + p_1^{n-1} p_1 = p_1^{n-1} a.$$

Therefore $p_1^{n-1} \subset a$. But

$$a = a(p_1^{n-1} + q_1^{n-1} + \cdots) = a p_1^{n-1} + a q_1^{n-1} + \cdots = a p_1^{n-1} + 0 + 0 + \cdots. \qquad \text{(T62)}.$$

Therefore $a \subset p_1^{n-1}$. Therefore $a = p_1^{n-1} = q_1 r_1 \cdots$.

Therefore $p_1^{n-1} q_1^{n-1} = 0$. $p_1^{n-1} \subset q_1$.

Therefore $q_1^{n-1} + r_1^{n-1} + \cdots \subset p_1$ (T42).

But $$p_1 = p_1(p_1^{n-1} + q_1^{n-1} + \cdots) = p_1 p_1^{n-1} + p_1(q_1^{n-1} + r_1^{n-1} + \cdots) = p_1(q_1^{n-1} + r_1^{n-1} + \cdots).$$

Therefore $p_1 \subset q_1^{n-1} + r_1^{n-1} + \cdots$. Therefore $p_1 = q_1^{n-1} + r_1^{n-1} + \cdots$.

T91. $$a \subset p_1^{n-1} :\rightarrow: (\exists i) \cdot a = 1^i p_1^{n-1}.$$

Proof. Hypothesis $:\rightarrow: a \subset q_1 r_1 \cdots \subset q_1, r_1, \cdots$.

Therefore $$a = (p_1 + a)(q_1 + a)(r_1 + a) \cdots = (p_1 + a) q_1 r_1 \cdots = (p_1 + 1^i) q_1 r_1 \cdots = p_1 q_1 \cdots + 1^i q_1 r_1 \cdots = 0 + 1^i p_1^{n-1}.$$

T92. $$a = 1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots.$$

Proof. $$a = a(p_1^{n-i} + q_1^{n-1} + \cdots) = a p_1^{n-1} + a q_1^{n-1} + \cdots = 1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots.$$

T93. $$1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots = (p_1 + 1^i)(q_1 + 1^j) \cdots.$$

Proof. Let $a = 1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots$. Then

$p_1 + a = p_1 + 1^i p_1^{n-1} = p_1 + 1^i$, by $1^j q_1^{n-1} \subset q_1^{n-1} \subset p_1$ and (T74).

Therefore $a = (p_1 + a)(q_1 + a) \cdots = (p_1 + 1^i)(q_1 + 1^j) \cdots$.

T94. $$(1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots)' = 1^{i+1} p_1^{n-1} + 1^{j+1} q_1^{n-1} + \cdots.$$

---

[26] This result can be trivially generalized: If p is a prime in K, p is prime in any sub-algebra of K containing p.

Proof. Let $a = 1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots$.

$$\begin{aligned} p_1 + a &= p_1 + 1^i. \\ p_1 + a' &= p_1 + (p_1 + a)' \\ &= p_1 + (p_1 + 1^i)' = p_1 + 1^{i+1}. \qquad \text{(T85)}. \end{aligned}$$

Similarly for $q_1 + a'$, etc.

Therefore
$$\begin{aligned} a' &= (p_1 + a')(q_1 + a') \cdots = (p_1 + 1^{i+1})(q_1 + 1^{j+1}) \cdots \\ &= 1^{i+1} p_1^{n-1} + 1^{j+1} q_1^{n-1} + \cdots. \end{aligned}$$

T95. *If*

$$a = 1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots = (p_1 + 1^i)(q_1 + 1^j) \cdots$$

and $f(x)$ *is definable in terms of* $+$ *and* $(')$, *then*

$$f(a) = f(1^i) p_1^{n-1} + f(1^j) q_1^{n-1} + \cdots = (p_1 + f(1^i))(q_1 + f(1^j)) \cdots.$$

Proof. If $f(x) \equiv x$, the theorem is obvious. If $f(x) \equiv b$, a constant, then

$$\begin{aligned} f(a) &= b(p_1^{n-1} + q_1^{n-1} + \cdots) = b p_1^{n-1} + b q_1^{n-1} + \cdots \\ &= f(1^i) p_1^{n-1} + f(1^i) q_1^{n-1} + \cdots. \end{aligned}$$

If $f(x) \equiv (g(x))'$, and the theorem is true for $g(x)$, then

$$\begin{aligned} p_1 + g(a) &= p_1 + g(1^i p_1^{n-1} + 1^j q_1^{n-1} + \cdots) \\ &= p_1 + g(1^i) p_1^{n-1} + g(1^j) q_1^{n-1} + \cdots = p_1 + g(1^i). \qquad \text{(T74)}. \end{aligned}$$

Therefore

$$\begin{aligned} p_1 + f(a) &= p_1 + g(a)' = p_1 + (p_1 + g(a))' \\ &= p_1 + (p_1 + g(1^i))' = p_1 + g(1^i)' = p_1 + f(1^i). \end{aligned}$$

Therefore

$$f(a) p_1^{n-1} = f(1^i) p_1^{n-1}.$$

Therefore

$$f(a) = f(a) p_1^{n-1} + f(a) q_1^{n-1} + \cdots = f(1^i) p_1^{n-1} + f(1^j) q_1^{n-1} + \cdots.$$

If $f(x) \equiv g(x) + h(x)$, and the theorem is true for $g(x)$ and $h(x)$ then

$$\begin{aligned} f(a) &= g(a) + h(a) = g(1^i) p_1^{n-1} + g(1^j) q_1^{n-1} + \cdots + h(1^i) p_1^{n-1} \\ &\quad + h(1^i) q_1^{n-1} + \cdots = (g(1^i) + h(1^i)) p_1^{n-1} + \cdots = f(1^i) p_1^{n-1} + \cdots. \\ p_1 &+ f(1^i) p_1^{n-1} + f(1^j) q_1^{n-1} + \cdots = p_1 + f(1^i) \quad \text{(by T74)}. \end{aligned}$$

Therefore

$$f(a) = (p_1 + f(a)) \cdots = (p_1 + f(1^i)) \cdots.$$

Therefore the theorem is true by induction.

We are now in a position to characterize the P_n of finite order precisely, and to prove our postulates consistent; they are also complete as soon as a postulate of the following form is added:

P12. *K contains exactly m elements.* (m finite).[27]

Let us consider all sequences of k numbers of the form

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \text{ where } 0 \leqq \alpha_i \leqq n-1. \qquad (i = 1, \cdots, k).$$

There are n^k such sequences. Define

$$\gamma = \alpha + \beta = (\gamma_1, \gamma_2, \cdots, \gamma_k),$$

where $\gamma_i = \min(\alpha_i, \beta_i)$, and $\delta = \alpha' = (\delta_1, \delta_2, \cdots, \delta_k)$, where $\delta_i =$ the least non-negative residue of $\alpha_i + 1$ modulo n.

It is easy to verify that these definitions of $+$ and $'$ yield a $P_n(n^k)$ (see Post, Webb for $k = 1$).[28]

Now take any P_n with k prime elements. Let $p, q, r, \cdots$ be its distinct prime elements. Form the correspondence

$$T(\alpha) = 1^{\alpha_1} p_1^{n-1} + 1^{\alpha_2} q_1^{n-1} + \cdots = (p_1 + 1^{\alpha_1})(q_1 + 1^{\alpha_2}) \cdots.$$

$T(\alpha)$ is a unique element of K. Conversely, if $T(\alpha) = T(\beta)$, $\alpha = \beta$ by the uniqueness of the decomposition proved in T86. Therefore T is one-to-one on $\{\alpha\}$ to K.[29]

$$\begin{aligned} T(\alpha') &= (T(\alpha))' \text{ by T94.} \\ T(\alpha) + T(\beta) &= (1^{\alpha_1} + 1^{\beta_1}) p_1^{n-1} + (1^{\alpha_2} + 1^{\beta_2}) q_1^{n-1} + \cdots \\ &= 1^{\min(\alpha_1,\beta_1)} p_1^{n-1} + 1^{\min(\alpha_2,\beta_2)} q_1^{n-1} + \cdots \\ &= T(\alpha + \beta). \end{aligned}$$

Therefore T is an isomorphism. But K was any P_n with k prime elements. Therefore every P_n with k prime elements is simply isomorphic with the algebra $\{\alpha\}$ defined above. These considerations yield:

T96. *If m is finite, any two $P_n(m)$ are simply isomorphic, and $m = n^k$, where k is the number of prime elements in each algebra. In particular, every $P_n(n^k)$ is isomorphic to the algebra defined above. The isomorphism may be obtained by pairing off the 1-elements*[30] *and the prime elements, and then*

---

[27] We show later that m must be a power of n.

[28] Cf. (7) in the bibliography below.

[29] We use here and below the fact that if $0 \leqq a, b \leqq n-1$, then $1^a + 1^b = 1^c$, where $c = \min(a, b)$. This follows almost immediately from T48.

[30] By 1-element we mean the element in each algebra represented by 1 in accordance with D4 and D14.

making two elements correspond to each other when they have corresponding representations of the type discussed in T92 (*or* T84).

This theorem contains in it the consistency and completeness properties announced above.

We can put the $P_n(n^k)$, $\{\alpha\}$, defined above, in a more suggestive form. Let $N = p_1^{n-1}p_2^{n-1}\cdots p_k^{n-1}$ where $p_1, p_2, \cdots, p_k$ are k distinct positive prime numbers[31] (in ordinary arithmetic), and K be the set of all positive divisors of N. Then any element a of K is of the form $a = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k} = p^\alpha$ symbolically. If we define the above operations on the exponents,

$$a \oplus b = p^{\alpha+\beta}, \qquad a' = p^{\alpha'},$$

then K, the set divisors of N, becomes a $P_n(n^k)$ (of course, $a \oplus b$ becomes "the greatest common divisor of a and b"). In the case $n = 2$, we get the Sheffer "Boolean numbers." It turns out that the p_i are the "prime elements" in this algebra, and the p_i^j are the "powers of primes"; indeed, these "Post numbers" were the starting point of the whole investigation embodied in this paper.

Some general remarks. In the last section we determined completely the structures of all possible Post algebras of finite order. The structure of Post algebras of infinite order is much more complicated and it seems as though a far from trivial extension of Stone's methods will be necessary to determine their structure as completely as he has determined the structure of infinite Boolean algebras. We intend to devote another paper to the study of infinite Post algebras. Among the results already obtained are:

There are P_n of any transfinite order with or without prime elements. This can be proved without assuming the continuum hypothesis.

Among the important unsolved problems in Post algebras are:

1) What are the necessary and sufficient conditions that an operation defined in terms of the Post operations be a group operation? 2) Are all Post algebras isomorphic to a subalgebra of the algebra of all functions on some given set A to the set of integers $0, 1, 2, \cdots, (n-1)$? (The operations on the functions are defined like the operations on the sequences discussed above.) 3) What are the lattice-theoretic characterizations of Post algebras? 4) Post constructed algebras by means of sequences of elements in a Boolean algebra.[32] (G. Birkhoff has communicated a similar construction

---

[31] Here p_i^j are the ordinary arithmetical powers of the p_i, not the Post operations in D9.

[32] Post (1), pp. 184-185.

to me privately.) Birkhoff conjectures that every Post algebra can be represented in this way. It is easy to show that this question is equivalent to problem 2) above. 5) What are the relations between Post algebras of Type n and combinatory topology modulo n? 6) Suppose we have a lattice with a 1-element. Define a power of a prime as an element a such that there is one and only one chain from a to 1. Then the finite Post algebras and ordinary cardinal arithmetic have the following proposition in common:

Every element can be factored uniquely into a product of powers of primes. (In arithmetic we use for "multiplication" the operation of finding the least common multiple.) Can purely lattice-theoretic conditions be found equivalent to this which would include as special cases Post algebras and arithmetic?

We wish to acknowledge the valuable help of the referee, Professor E. L. Post, in the preparation of this paper. He not only made contributions to the actual material of the paper (see pp. 176-178) but also made suggestions concerning style and exposition.

UNIVERSITY OF PENNSYLVANIA.

BIBLIOGRAPHY.

(1) E. L. Post, "Introduction to a general theory of elementary propositions," *American Journal of Mathematics*, vol. 43 (1921), pp. 163-185, esp. pp. 180-185.

(2) D. L. Webb, "The algebra of n-valued logic," *C. R. Soc. Sci. Varsovie*, vol. 29 (1937), pp. 153-168.

(3) ———, "Generation of any n-valued logic by one binary operation," *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 252-254.

(4) J. Lukasiewicz and A. Tarski, "Untersuchungen über den Aussagenkalkül," *C. R. Soc. Sci. Lett. Varsovie*, Classe III, vol. 23 (1930), pp. 1-21.

(5) B. A. Bernstein, "Modular representations of finite algebras," *Proceedings of the International Mathematical Congress*, Toronto, vol. I (1924), pp. 207-216.

(6) Orrin Fink, Jr., "New algebras of logic," *The American Mathematical Monthly*, vol. 45 (1938), pp. 210-219.

(7) R. Vaidayanathaswamy, "Quasi-Boolean algebras and many-valued logics," *Proceedings of the Indian Academy of Sciences*, vol. 8, sec. A (1938), pp. 165-170.

(8) J. Barkley Rosser, "The introduction of quantification into a three-valued logic," *The Journal of Unified Science (Erkenntnis)*, vol. 9 (1939), pp. 136-141.

ON SUMS OF POSITIVE INTEGRAL k-TH POWERS.*

By H. DAVENPORT.

In a previous paper [1] with the same title, the author has established certain lower bounds for $N_s^{(k)}(n)$, the number of numbers less than n that are representable as the sum of s positive integral k-th powers. The object of the present paper is to prove more precise results of the same type. A knowledge of the previous paper is not assumed.

In the paper which follows this, these results are applied to Waring's problem to prove that $G(5) \leqq 23$, $G(6) \leqq 36$.

All small Latin letters (with or without suffixes) denote positive integers, and $k \geqq 3$. ϵ denotes an arbitrarily small positive number. The constants implied by the symbol O depend only on k and ϵ.

THEOREM 1. *Suppose that*

(a) $u_1 < u_2 < \cdots$; (b) $0 < \delta < 1$, $\lambda = 1 - (1-\delta)/k$; (c) $l \leqq k-2$.

Then the number S of solutions of

$$x^k + u_i = y^k + u_j \tag{1}$$

subject to

$$P < x,\ y < 2P, \qquad u_i, u_j < P^{k\lambda} \tag{2}$$

satisfies

$$S = O(PU + P^{1+\delta+\epsilon}U\{P^{-2} + P^{-\delta-l-1}U\}^{1/2^l}), \tag{3}$$

where U denotes the number of u's less than $P^{k\lambda}$.

Proof. The number of solutions with $x = y$ is $O(PU)$. Hence

$$S = O(PU) + O(M_1), \tag{4}$$

where M_1 denotes the number of solutions with $y > x$. Writing $y = x + t$, (1) becomes

$$\Delta_t(x^k) + u_j = u_i, \tag{5}$$

where $\Delta_t(F(x)) = F(x+t) - F(x)$. Plainly

---

* Received March 28, 1940. This paper was originally submitted for publication to the Polish journal *Acta Arithmetica* in March 1939, but the number which was to have contained it never appeared.

[1] *Proceedings of the Royal Society* (A), vol. 170 (1939), pp. 293-299. See also a paper by Davenport and Erdös, *Annals of Mathematics*, vol. 40 (1939), pp. 533-536.

$$tP^{k-1} < tx^{k-1} < \Delta_t(x^k) < u_i < P^{k\lambda},$$

whence

(6) $$t < P^{\delta}.$$

In the remainder of the proof, variables denoted by x (and similarly x'), u_i, u_j (and similarly $u_{j'}$), t are subject to (2), (6).

Let $M(t, u_i)$ denote the number of solutions of (5) for given t, u_i. By Cauchy's inequality,

$$M_1 \leqq \sum_{t,u_i} M(t, u_i) \leqq (\sum_{t,u_i} 1)^{1/2}(\sum_{t,u_i} M^2(t, u_i))^{1/2} < (P^{\delta}U)^{1/2}(M'_1)^{1/2},$$

where M'_1 denotes the number of solutions of

(7) $$\Delta_t(x^k) + u_j = \Delta_t(x'^k) + u_{j'} = u_i.$$

The number of solutions of (7) with $x' = x$ is precisely M_1. In the solutions with $x' > x$, write $x' = x + t_1$. (7) implies

(8) $$\Delta_{t,t_1}(x^k) + u_{j'} = u_j,$$

where $\Delta_{t,t_1} = \Delta_{t_1}(\Delta_t)$, and t_1 is subject to $t_1 < P$. Thus

$$M'_1 \leqq M_1 + 2M_2 = O\ (\max(M_1, M_2)),$$

where M_2 denotes the number of solutions of (8). Hence

$$M_1 = O((P^{\delta}U)^{1/2} \max(M_1^{1/2}, M_2^{1/2})),$$

which implies

(9) $$M_1 = O(P^{\delta}U) + O((P^{\delta}UM_2)^{1/2}).$$

We now generalize the above argument. Let M_h denote the number of solutions of

(10) $$\Delta_{t,t_1,\ldots,t_{h-1}}(x^k) + u_j = u_i,$$

where

(11) $$t_1, t_2, \cdots < P.$$

Applying Cauchy's inequality in the same way as before, we obtain

$$M_h < (P^{\delta+h-1}U)^{1/2}(M'_h)^{1/2}; \qquad M'_h \leqq M_h + 2M_{h+1}.$$

Hence

(12) $$M_h = O(P^{\delta+h-1}U) + O((P^{\delta+h-1}UM_{h+1})^{1/2}).$$

It follows that

(13) $$M_1 = O(P^{1+\delta}U\{P^{-2} + M_{l+1}P^{-\delta-l-1}U^{-1}\}^{1/2^l}).$$

For, if $l = 1$, this reduces to (9); and if (13) holds for $l = h - 1$ it holds for $l = h$, since

$$\begin{aligned}\{P^{-2}+M_hP^{-\delta-h}U^{-1}\}^{1/2^{h-1}} &= O\{P^{-4}+M_h{}^2P^{-2\delta-2h}U^{-2}\}^{1/2^h}\\ &= O\{P^{-4}+P^{-2\delta-2h}U^{-2}(P^{2\delta+2h-2}U^2+P^{\delta+h-1}UM_{h+1})\}^{1/2^h}\\ &= O\{P^{-2}+M_{h+1}P^{-\delta-h-1}U^{-1}\}^{1/2^h}.\end{aligned}$$

We now prove that

$$M_{l+1}=O(P^{\epsilon}U^2). \tag{14}$$

Given u_i, u_j in (10), where $h=l+1$, the variables $t, t_1, \cdots, t_l$ are determined with only $O(P^{\epsilon})$ possibilities as factors of u_i-u_j. Since, by (c), $l+1\leqq k-1$, the remaining factor in $\Delta_{t,t_1,\ldots,t_l}(x^k)$ then determines x uniquely.

The assertion of the theorem follows from (4), (13), (14).

THEOREM 2. *If $N_s^{(k)}(n)>n^{\alpha}$ for all large n, where $1/k<\alpha<1$, then*

$$N_{s+1}^{(k)}(n)>n^{\beta-\epsilon} \tag{15}$$

for $n>n_0(\epsilon)$, where

$$\beta=\max_{h\leqq k-2}\frac{1}{k}\left\{1+\frac{(2^h-1)(k-1)+h+1}{2^h-1+\alpha}\alpha\right\}. \tag{16}$$

Proof. Let l be the value of h for which the maximum is attained. Let

$$\delta=\{l+1-(k-1)\alpha\}/(2^l-1+\alpha). \tag{17}$$

Then δ is also a maximum for this value of l and so $\delta>0$. Also

$$l+1-(k-1)\alpha\leqq 2^l-k\alpha+\alpha<2^l-1+\alpha,\quad \delta<1.$$

Let

$$\lambda=1-\frac{1}{k}+\frac{\delta}{k}=\frac{(2^l-1)(k-1)+l+1}{k(2^l-1+\alpha)}.$$

By (16),

$$k\beta=1+k\lambda\alpha. \tag{18}$$

We observe for future reference that

$$\delta+k\lambda\alpha/2^l=(\delta+l+1)/2^l. \tag{19}$$

Since l is the integer for which δ (defined by (17)) is a maximum, we have, provided $l>1$,

$$\frac{l-(k-1)\alpha}{2^{l-1}-1+\alpha}\leqq\frac{l+1-(k-1)\alpha}{2^l-1+\alpha},$$

whence

$$(l+1-(k-1)\alpha)2^{l-1}\leqq 2^l-1+\alpha,\quad\text{i. e.}$$

$$\delta\leqq 1/2^{l-1}. \tag{20}$$

Obviously this holds also when $l=1$.

Let $P = n^{1/k}/3$. Let $u_1 < u_2 < \cdots < u_U < P^{k\lambda}$ denote the numbers less than $P^{k\lambda}$ which are representable as the sum of s positive integral k-th powers. By hypothesis,

$$U > P^{k\lambda\alpha}. \tag{21}$$

The conditions of Theorem 1 are satisfied. Let $r(m)$ denote the number of representations of m as $x^k + u_i$, where $P < x < 2P$. By Theorem 1,

$$\sum_m r^2(m) = O(PU + P^{1+\delta+\epsilon}U\{P^{-2} + P^{-\delta-l-1}U\}^{1/2^l}).$$

Also, plainly, $\sum_m r(m) > (P-2)U$. By (21), $PU = O(P^{1-k\lambda\alpha}U^2)$. By (20), (21),

$$P^{1+\delta+\epsilon}U(P^{-2})^{1/2^l} \leqq P^{1+\epsilon}U = O(P^{1-k\lambda\alpha+\epsilon}U^2).$$

By (19), (21),

$$P^{1+\delta+\epsilon}U(P^{-\delta-l-1}U)^{1/2^l} = O(P^{1-k\lambda\alpha+\epsilon}U^2).$$

Hence

$$P^{1+k\lambda\alpha-\epsilon}\sum_m r^2(m) = O((\sum_m r(m))^2).$$

Using Cauchy's inequality, we have

$$N_{s+1}^{(k)}(n) \geqq \sum_{\substack{m \\ r(m)>0}} 1 \geqq (\sum_m r(m))^2/\sum_m r^2(m) > P^{1+k\lambda\alpha-\epsilon}/O(1) > P^{1+k\lambda\alpha-2\epsilon}$$
$$= P^{k\beta-2\epsilon} \text{ by (18) } > n^{\beta-\epsilon},$$

since $2\epsilon/k < \epsilon$. This proves Theorem 2.

THEOREM 3. *$N_s^{(4)}(n) > n^{\alpha_s-\epsilon}$ for $n > n_0(\epsilon)$, where*

$$\alpha_2 = 1/2, \qquad \alpha_3 = 19/28, \qquad \alpha_4 = 331/412, \qquad \alpha_5 = 5539/6268.$$

Proof. The first result is well known, and follows from the fact that the number of representations of an integer m as the sum of two fourth powers is $O(m^\epsilon)$.

By Theorem 2, taking $h = 2$ in (16), an admissible value for α_s $(s > 2)$ is

$$\alpha_s = \frac{1}{4}\left\{1 + \frac{3\cdot 3 + 2 + 1}{3 + \alpha_{s-1}}\alpha_{s-1}\right\} = \frac{3 + 13\alpha_{s-1}}{4(3 + \alpha_{s-1})}.$$

The results follow by repeated application of this formula.[2]

We now prove a theorem, analogous to Theorem 1, which yields more precise results when $k = 5$ or 6.

---

[2] These results were proved and applied to Waring's problem in a paper in the *Annals of Mathematics*, vol. 40 (1939), pp. 731-747.

THEOREM 4. *Assume hypotheses* (a) *and* (b) *of Theorem* 1, *but suppose that* (c) *either* $l = k - 3$, *or* $l \leqq k - 4$ *and* $k - l$ *is even. Then the number* S *of solutions of*

$$x^k + p^k u_i = y^k + p^k u_j, \tag{22}$$

subject to

$$p \text{ prime}, \qquad p \equiv -1 \pmod{k}, \qquad 2^{-1/k} P^{1-\lambda} < p < P^{1-\lambda}, \tag{23}$$

$$P < x,\ y < 2P, \qquad u_i, u_j < P^{k\lambda}, \qquad p \nmid xy, \tag{24}$$

satisfies

$$S = O(PU\Pi + P^{1+\delta+\epsilon} U\Pi \{P^{-2} + P^{-\delta-l-1}\Pi^{-1}U\}^{1/2^l}), \tag{25}$$

where U *denotes the number of* u*'s less than* $P^{k\lambda}$, *and* Π *denotes the number of primes* p *satisfying* (23).

Proof. The number of solutions with $x = y$ is $O(PU\Pi)$.

Since $p > k$ and $p \equiv -1 \pmod{k}$, we have $(k, \phi(p^k)) = (k, 2)$. Hence, if $p \nmid x$, the congruence $y^k \equiv x^k \pmod{p^k}$ implies $y \equiv \pm x \pmod{p^k}$, where the negative sign is possible only when k is even. Thus, in the solutions of (22) with $y > x$, either $y = x + tp^k$ or $y = -x + tp^k$, the latter possibility arising only if k is even. In the first case, we write $x = \xi$, in the second case $x = -\xi$. In either case, $2\xi + tp^k > 0$. Then

$$S = O(PU\Pi) + O(M_1), \tag{26}$$

where M_1 denotes the number of solutions of

$$\xi^k + p^k u_i = (\xi + tp^k)^k + p^k u_j,$$

that is, of

$$p^{-k} \Delta_{tp^k}(\xi^k) + u_j = u_i, \tag{27}$$

where ξ is an integral variable, and the conditions (24) are replaced by

$$P < |\xi| < 2P, \qquad u_i, u_j < P^{k\lambda}, \qquad p \nmid \xi. \tag{28}$$

Since $tp^k = y - \xi$, we have

$$tP^{k(1-\lambda)} < 8P.$$

Hence t satisfies

$$t < 8P^\delta, \qquad 2\xi + tp^k > 0. \tag{29}$$

In the remainder of the proof, variables denoted by ξ (and similarly ξ'), u_i, u_j (and similarly $u_{j'}$), t, p are subject to (23), (28), (29).

Let $M(t, u_i, p)$ denote the number of solutions of (27) for given t, u_i, p. By Cauchy's inequality,

$$\begin{aligned} M_1 \leqq \sum_{t,u_i,p} M(t, u_i, p) &\leqq \Big(\sum_{t,u_i,p} 1\Big)^{1/2} \Big(\sum_{t,u_i,p} M^2(t, u_i, p)\Big)^{1/2} \\ &< (8P^\delta U\Pi)^{1/2} (M'_1)^{1/2}, \end{aligned}$$

where M'_1 denotes the number of solutions of

$$p^{-k}\Delta_{tp^k}(\xi^k) + u_j = p^{-k}\Delta_{tp^k}(\xi'^k) + u_{j'} = u_i.$$

The number of solutions with $\xi' = \xi$ is precisely M_1. Hence

$$M_1 = O(P^\delta U\Pi) + O((P^\delta U\Pi M_2)^{\frac12}),$$

where M_2 denotes the number of solutions of

$$p^{-k}\Delta_{tp^k, t_1}(\xi^k) + u_{j'} = u_j, \qquad t_1 < 4P.$$

The argument generalizes exactly as in the proof of Theorem 1. Denoting by M_{l+1} the number of solutions of

$$p^{-k}\Delta_{tp^k, t_1, \ldots, t_l}(\xi^k) + u_j = u_i, \qquad t_1, t_2, \cdots < 4P, \tag{30}$$

we obtain

$$M_1 = O(P^{1+\delta}U\Pi\{P^{-2} + P^{-\delta-l-1}U^{-1}\Pi^{-1}M_{l+1}\}^{1/2^l}). \tag{31}$$

We now prove that

$$M_{l+1} = O(P^\epsilon U^2). \tag{32}$$

We note first the identity

$$\Delta_{w_1, \ldots, w_r}(\xi^k) = 2^{-k} \sum_{\theta_1, \ldots, \theta_r} \theta_1 \cdots \theta_r(\zeta + \theta_1 w_1 + \cdots + \theta_r w_r)^k, \tag{33}$$

where $\zeta = 2\xi + w_1 + \cdots + w_r$, and each of $\theta_1, \cdots, \theta_r$ takes the values 1 and -1. This follows at once from the definition of the symbol Δ. In the expansion of $(\zeta + \theta_1 w_1 + \cdots + \theta_r w_r)^k$ on the right of (33), the only terms which remain after summation over $\theta_1, \cdots, \theta_r$ are those in which each of $w_1, \cdots, w_r$ occurs to an odd power. Hence, firstly, if $k - r$ is odd,

$$\Delta_{w_1, \ldots, w_r}(\xi^k) = 2^{-k}\zeta w_1 \cdots w_r F(\zeta^2, w_1^3, \cdots, w_r^2), \tag{34}$$

where F is a homogeneous polynomial of degree $\frac12(k - r - 1)$ whose coefficients are positive integers depending only on k and r. Secondly, if $k - r = 2$,

$$\begin{aligned} \Delta_{w_1, \ldots, w_r}(\xi^k) &= 2^{-k}2^r w_1 \cdots w_r(k!\,\zeta^2/2! + k!\,(w_1^2 + \cdots + w_r^2)/3!) \\ &= k!\,w_1 \cdots w_r(3\zeta^2 + w_1^2 + \cdots + w_r^2)/24. \end{aligned} \tag{35}$$

We apply (34) and (35) to prove (32).

Case 1. $l \leq k - 4$, $k - l$ even. By (34) with $r = l + 1$, the equation (30) has the form

$$\zeta t t_1 \cdots t_l F(\zeta^2, (tp^k)^2, t_1^2, \cdots, t_l^2) = 2^k(u_i - u_j), \tag{36}$$

where

$$\zeta = 2\xi + tp^k + t_1 + \cdots + t_l. \tag{37}$$

By the second half of (29), $\zeta > 0$. Given u_i, u_j with $u_i - u_j > 0$ (since the left-hand side of (36) is positive), there are only $O(P^\epsilon)$ possibilities for all of $\zeta, t, t_1, \cdots, t_l$ as factors of $2^k(u_i - u_j)$. For any such set of values, (36) determines p uniquely, and (37) then determines ξ uniquely. This proves (32) in this case.

Case 2. $l = k - 3$. By (35) with $r = l + 1 = k - 2$, the equation (30) has the form

$$(38) \qquad k!\, t t_1 \cdots t_l (3\zeta^2 + (tp^k)^2 + t_1^2 + \cdots + t_l^2) = 24(u_i - u_j).$$

Given u_i, u_j with $u_i - u_j > 0$ (since the left-hand side of (38) is positive), there are only $O(P^{\epsilon/2})$ possible sets of values for $t, t_1, \cdots, t_l$ as factors of $24(u_i - u_j)$. For any such set of values, (38) determines ζ and p with only $O(P^{\epsilon/2})$ possibilities, since the number of representations of an integer m as $3\zeta^2 + q^2$ is $O(m^\epsilon)$. This proves (32) in this case.

The assertion of the theorem follows from (26), (31), (32).

THEOREM 5. *$N_s^{(5)}(n) > n^{\alpha_s - \epsilon}$ for $n > n_0(\epsilon)$, where $\alpha_2 = 2/5$ and*

$$(39) \qquad \alpha_s = \frac{16 + 85\alpha_{s-1}}{5(16 + 5\alpha_{s-1})} \quad \text{if} \quad 2/5 \leqq \alpha_{s-1} \leqq 3/5,$$

$$(40) \qquad \alpha_s = \frac{7 + 33\alpha_{s-1}}{5(7 + \alpha_{s-1})} \quad \text{if} \quad 3/5 \leqq \alpha_{s-1} \leqq 1.$$

Note. The values of α_s given by these formulae are $\alpha_3 = 5/9$, $\alpha_4 = 569/845$, and thereafter by (40).

The result is well known for $s = 2$. We therefore assume that $N_{s-1}^{(5)}(n) > n^{\alpha_{s-1} - \epsilon}$ for $n > n_0(\epsilon)$, and deduce the corresponding result for s, where α_s is given by (39) or (40).

Proof of (39). Let

$$(41) \qquad \delta = (16 - 20\alpha)/(16 + 5\alpha), \quad \lambda = (4 + \delta)/5 = 16/(16 + 5\alpha),$$

where, for brevity, α is written for α_{s-1}. Plainly $\delta > 0$. Also

$$(42) \qquad \delta < [16 - 20(2/5)]/16 = 1/2.$$

We observe for future reference that

$$(43) \qquad \delta + 5\lambda\alpha/4 = (\delta + 3)/4 + (1 - \lambda)/4.$$

Let $P = n^{1/5}/3$. Let $u_1 < u_2 < \cdots < u_U < P^{5\lambda}$ be the numbers less than $P^{5\lambda}$ representable as the sum of $s - 1$ positive integral fifth powers. By hypothesis,

$$(44) \qquad U > P^{5\lambda\alpha - \epsilon}.$$

Let $r(m)$ denote the number of representations of an integer m as $x^5 + p^5 u_i$, where $P < x < 2P$, p prime, $p \equiv -1 \pmod 5$, $p \nmid x$, $2^{-1/5}P^{1-\lambda} < p < P^{1-\lambda}$. Plainly

$$\sum_m r(m) > PU\Pi/2, \tag{45}$$

where Π denotes the number of primes $p \equiv -1 \pmod 5$ satisfying $2^{-1/5}P^{1-\lambda} < p < P^{1-\lambda}$. It is well known that

$$\Pi > CP^{1-\lambda}/\log P, \tag{46}$$

for large P, where C is a positive absolute constant.

Let $r(p, m)$ denote the number of representations of m in the above form for given p. Using Cauchy's inequality, we have

$$\begin{aligned}\sum_m r^2(m) &= \sum_m \Bigl(\sum_p r(p,m)\Bigr)^2 \leqq \sum_m \Bigl(\sum_p 1\Bigr)\Bigl(\sum_p r^2(p,m)\Bigr)\\ &= \Pi \sum_m \sum_p r^2(p,m) = \Pi S,\end{aligned}$$

where S denotes the number of solutions of (22) subject to (23), (24), with $k = 5$. By Theorem 4 with $k = 5$, $l = k - 3 = 2$,

$$S = O(PU\Pi + P^{1+\delta+\epsilon}U\Pi\{P^{-2} + P^{-\delta-3}\Pi^{-1}U\}^{1/4}).$$

By (44),

$$PU\Pi = O(P^{1-5\lambda\alpha+\epsilon}U^2\Pi).$$

By (42), (44),

$$P^{1+\delta+\epsilon}U\Pi(P^{-2})^{1/4} < P^{1+\epsilon}U\Pi = O(P^{1-5\lambda\alpha+2\epsilon}U^2\Pi).$$

By (43), (44), (46),

$$P^{1+\delta+\epsilon}U\Pi(P^{-\delta-3}\Pi^{-1}U)^{1/4} = O(P^{1-5\lambda\alpha+2\epsilon}U^2\Pi(\log P)^{1/4}) = O(P^{1-5\lambda\alpha+3\epsilon}U^2\Pi).$$

Hence

$$P^{1+5\lambda\alpha-3\epsilon} \sum_m r^2(m) = O\Bigl(\Bigl(\sum_m r(m)\Bigr)^2\Bigr). \tag{47}$$

We have

$$\begin{aligned}N_s^{(5)}(n) \geqq \sum_{\substack{m\\ r(m)>0}} 1 \geqq \Bigl(\sum_m r(m)\Bigr)^2 \Big/ \sum_m r^2(m) &> P^{1+5\lambda\alpha-3\epsilon}/O(1) \text{ by (47)}\\ &> P^{1+5\lambda\alpha-4\epsilon} > n^{\beta-\epsilon},\end{aligned}$$

where

$$\beta = (1 + 5\lambda\alpha)/5 = ((16/5) + 17\alpha)/(16 + 5\alpha).$$

Proof of (40). By Theorem 2 with $k = 5$, $h = 3$, an admissible value for α_s is

$$\alpha_s = [1 + \{(7 \cdot 4 + 4)\alpha_{s-1}/(7 + \alpha_{s-1})\}]/5 = (7 + 33\alpha_{s-1})/5(7 + \alpha_{s-1}).$$

THEOREM 6. *$N_s^{(6)}(n) > n^{\alpha_s - \epsilon}$ for $n > n_0(\epsilon)$, where $\alpha_2 = 2/6$ and*

(48) $\alpha_s = (19 + 120\alpha_{s-1})/6(19 + 6\alpha_{s-1})$ *if* $1/3 \leqq \alpha_{s-1} \leqq 19/42$,

(49) $\alpha_s = (43 + 246\alpha_{s-1})/6(43 + 6\alpha_{s-1})$ *if* $19/42 \leqq \alpha_{s-1} \leqq 2/3$,

(50) $\alpha_s = (15 + 81\alpha_{s-1})/6(15 + \alpha_{s-1})$ *if* $2/3 \leqq \alpha_{s-1} < 1$.

Note. The values of α_s given by these formulae are

$\alpha_3 = 59/126$, $\alpha_4 = 1661/2886$, $\alpha_5 = 5549/8379$, $\alpha_6 = 575117/787182$, and thereafter by (50).

The result is well known for $s = 2$. We therefore assume that $N^{(6)}_{s-1}(n) > n^{\alpha_{s-1}-\epsilon}$ for $n > n_0(\epsilon)$, and deduce the corresponding result for s, where α_s is given by (48) or (49) or (50).

Proof of (48). Let

$$\delta = (19 - 30\alpha)/(19 + 6\alpha), \qquad \lambda = (5 + \delta)/6 = 19/(19 + 6\alpha),$$

where α is written for α_{s-1}. Plainly $\delta > 0$. Also

$$\delta < 1 - (10/19) < 1/2.$$

We observe that

$$\delta + 6\lambda\alpha/4 = (\delta + 3)/4 + (1 - \lambda)/4.$$

The rest of the proof is almost the same as in the proof of (39). The inequality (44) is replaced by

$$U > P^{6\lambda\alpha-\epsilon}.$$

We appeal to Theorem 4 with $k = 6$, $l = 2$, which is permissible, since $k - l$ is then even. We obtain (47) with the exponent $1 + 6\lambda\alpha - 3\epsilon$ in place of $1 + 5\lambda\alpha - 3\epsilon$, and finally

$$N_s^{(6)}(n) > n^{\beta-\epsilon},$$

where

$$\beta = (1 + 6\lambda\alpha)/6 = (19 + 120\alpha)/6(19 + 6\alpha).$$

Proof of (49). Let

$$\delta = (25 - 30\alpha)/(43 + 6\alpha), \qquad \lambda = (5 + \delta)/6 = 40/(43 + 6\alpha),$$

where α is written for α_{s-1}. Plainly $\delta > 0$. Also

(51) $$\delta \leqq [25 - 30(19/42)]/[43 + 6(19/42)] = 1/4.$$

We observe for future reference that

(52) $$\delta + 6\lambda\alpha/8 = (\delta + 4)/8 + (1 - \lambda)/8.$$

Let $P = n^{1/6}/3$. Let $u_1 < u_2 < \cdots < u_U < P^{6\lambda}$ be the numbers less than $P^{6\lambda}$ representable as the sum of $s-1$ sixth powers. By hypothesis,

$$U > P^{6\lambda\alpha-\epsilon}. \tag{53}$$

Defining $r(m)$ and $r(p, m)$ in the same way as before, we have

$$\sum_m r(m) > PU\Pi/2$$

where Π satisfies (46); and

$$\sum_m r^2(m) \leqq \Pi S,$$

where S denotes the number of solutions of (22) subject to (23), (24) with $k = 6$. By Theorem 4 with $k = 6$, $l = k - 3 = 3$,

$$S = O(PU\Pi + P^{1+\delta+\epsilon}U\Pi\{P^{-2} + P^{-\delta-4}\Pi^{-1}U\}^{1/8}).$$

By (53),

$$PU\Pi = O(P^{1-6\lambda\alpha+\epsilon}U^2\Pi).$$

By (51), (53),

$$P^{1+\delta+\epsilon}U\Pi(P^{-2})^{1/8} \leqq P^{1+\epsilon}U\Pi = O(P^{1-6\lambda\alpha+2\epsilon}U^2\Pi).$$

By (52), (53), (46),

$$P^{1+\delta+\epsilon}U\Pi(P^{-\delta-4}\Pi^{-1}U)^{1/8} = O(P^{1-6\lambda\alpha+3\epsilon}U^2\Pi).$$

Hence

$$P^{1+6\lambda\alpha-3\epsilon}\sum_m r^2(m) = O((\sum_m r(m))^2).$$

It follows as before that $N_s^{(6)}(n) > n^{\beta-\epsilon}$, where

$$\beta = (1 + 6\lambda\alpha)/6 = (43 + 246\alpha)/6(43 + 6\alpha).$$

Proof of (50). By Theorem 2 with $k = 6$, $h = 4$, an admissible value for α_s is

$$\alpha_s = [1 + (15 \cdot 5 + 5)\alpha_{s-1}/(15 + \alpha_{s-1})]/6 = (15 + 81\alpha_{s-1})/6(15 + \alpha_{s-1}).$$

THE UNIVERSITY,
MANCHESTER, ENGLAND.

ON WARING'S PROBLEM FOR FIFTH AND SIXTH POWERS.[*]

By H. DAVENPORT.

Introduction. Let $G(k)$ denote the least value of s for which it is true that every sufficiently large positive integer is representable as the sum of s positive integral k-th powers. In a recent paper,[1] I have proved (*inter alia*) that $G(4) = 16$. In the present paper, it will be proved that

$$G(5) \leqq 23, \quad G(6) \leqq 36. \tag{1}$$

Up to now, the best known results (due to Hua and Estermann respectively [2]) have been $G(5) \leqq 28$, $G(6) \leqq 42$.

The necessary new material for the proof is provided in the paper preceding this, which will be referred to as II. The proof as a whole follows the same lines as for fourth powers, and many of the lemmas are taken from I and from Landau.[3]

The fundamental lemmas. Throughout the paper, k is either 5 or 6. We define

$$s_1 = 8 \text{ if } k = 5; \ = 13 \text{ if } k = 6, \qquad s_2 = 7 \text{ if } k = 5; \ = 10 \text{ if } k = 6.$$

Let $N_s^{(k)}(n)$ denote the number of numbers less than n that are representable as the sum of s positive integral k-th powers.

LEMMA 1. *For any* $\epsilon > 0$ *we have* $N_s^{(k)}(n) > n^{\mu_s^{(k)} - \epsilon}$ *for* $n > n_0(\epsilon)$, *where*

$$\mu_4^{(5)} = 569/845, \qquad \mu_s^{(5)} = (7 + 33\mu_{s-1}^{(5)})/5(7 + \mu_{s-1}^{(5)}) \qquad (s \geqq 5),$$
$$\mu_6^{(6)} = 575117/787182, \qquad \mu_s^{(6)} = (15 + 81\mu_{s-1}^{(6)})/6(15 + \mu_{s-1}^{(6)}) \qquad (s \geqq 7).$$

Proof. Theorems 5, 6 of II.

LEMMA 2. $k\mu_{s_1}^{(k)} + s_2/2^{k-1} > k + 10^{-4}$.

Proof. By calculation from Lemma 1, we have

$$\mu_8^{(5)} = 6913439/7576115 > (5 - (7/16) + 10^{-4})/5,$$
$$\mu_{13}^{(6)} = \frac{24040\,980\,990\,984\,981}{25335\,323\,032\,000\,606} > \frac{1}{6}\left(6 - \frac{10}{32} + 10^{-4}\right).$$

---

[*] Received March 28, 1941.

[1] *Annals of Mathematics*, vol. 40 (1939), pp. 731-747. This will be referred to as I.

[2] Hua, *Proceedings of the London Mathematical Society*, ser. 2, vol. 45 (1937), pp. 144-160; Estermann, *Acta Arithmetica*, vol. 2 (1937), pp. 197-211.

[3] *Vorlesungen über Zahlentheorie* (Leipzig, 1927), volume 1.

From now onwards, s is defined by

$$s = 2s_1 + s_2 = 23 \text{ if } k = 5; \ = 36 \text{ if } k = 6.$$

Notation. Throughout the paper, all small Latin letters, with or without suffixes, denote positive integers. δ is a small positive number, fixed throughout, and ϵ is an arbitrarily small positive number, not necessarily the same at each occurrence. $C_1, C_2, \cdots$ denote positive absolute constants.

Let N be the large number which is to be represented as the sum of s positive integral k-th powers, and let

$$(2) \qquad P = [(N/100)^{1/k}].$$

Let $u_1, u_2, \cdots, u_U$ be the (different) numbers less than P^k that are representable as the sum of s_1 positive integral k-th powers. By Lemma 1,

$$(3) \qquad U > P^{k\beta_{s_1}{}^{(k)}-\epsilon}.$$

For any real α, let

$$T(\alpha) = \sum_{x=P}^{2P} e(\alpha x^k); \qquad U(\alpha) = \sum_{h=1}^{U} e(\alpha u_h),$$

where $e(A)$ is an abbreviation for $e^{2\pi i A}$. Let

$$T^{s_2}(\alpha) U^2(\alpha) = \sum_m r_s(m) e(m\alpha).$$

Then $r_s(m)$ does not exceed the number of representations of m as the sum of $s = 2s_1 + s_2$ positive integral k-th powers, and in order to prove (1) it suffices to prove that $r_s(N) > 0$. We have

$$(4) \qquad r_s(N) = \int T^{s_2}(\alpha) U^2(\alpha) e(-N\alpha)\, d\alpha,$$

where the range of integration is any interval of length 1.

Throughout the paper, a and q are subject to $a \leqq q$, $(a, q) = 1$. Let

$$S_{a,q} = \sum_{x=1}^{q} e_q(ax^k) \qquad (e_q(A) = e(A/q)).$$

For any real β, let

$$I(\beta) = \sum_{n=P^k}^{(2P)^k} n^{-1+(1/k)} e(\beta n)/k,$$

and let

$$T^*(\alpha, a, q) = q^{-1} S_{a,q} I(\alpha - a/q).$$

Approximations to $T(\alpha)$.

LEMMA 3. $S_{a,q} = O(q^{1-(1/k)})$.

Proof. Landau, Satz 315.

LEMMA 4. *If* $|\beta| \leqq 1/2$, *then*

$$I(\beta) = O(\min(P, P^{1-k}|\beta|^{-1})).$$

Proof. Lemma 4 of I.

LEMMA 5. *If* $\alpha = a/q + \beta$, *where* $|\beta| \leqq 1/2$, *then*

$$T^*(\alpha, a, q) = O(q^{-1/k}\min(P, P^{1-k}|\beta|^{-1})).$$

Proof. Lemmas 3, 4.

LEMMA 6. *If* $\alpha = a/q + \beta$, *where* $q \leqq P^{1-\delta}$ *and* $\beta = O(q^{-1}P^{1-k-\delta})$, *then*

$$T(\alpha) - T^*(\alpha, a, q) = O(q^{(3/4)+\epsilon}); \qquad T(\alpha) = O(q^{-1/k}\min(P, P^{1-k}|\beta|^{-1})).$$

Proof. Lemmas 8, 9 of I.

LEMMA 7. *If* $P^{1-\delta} \leqq q \leqq P^{k-1+\delta}$, *and* $\beta = O(q^{-1}P^{1-k-\delta})$, *then*

$$T(\alpha) = O(P^{1-\kappa+\delta}),$$

where $\kappa = 1/2^{k-1}$.

Proof. Lemma 11 of I.

The Farey dissection. Let $Q = [P^{k-1+\delta}]$. For $q \leqq P^{1/2}$, let $\mathfrak{M}_{a,q}$ denote the interval

$$|\alpha - a/q| \leqq 1/qQ.$$

Two intervals $\mathfrak{M}_{a_1,q_1}$, $\mathfrak{M}_{a_2,q_2}$ corresponding to different pairs a, q do not overlap, since

$$|(a_1/q_1) - (a_2/q_2)| \geqq 1/q_1q_2 > 2P^{(3/2)-k-\delta}/q_1q_2 \geqq Q^{-1}[(1/q_1) + (1/q_2)].$$

The hypotheses of Lemmas 5, 6 are satisfied if α is in any $\mathfrak{M}_{a,q}$.

Denote by $\mathfrak{m}$ all points of the interval

$$1/Q < \alpha < 1 + 1/Q \tag{5}$$

which do not belong to any $\mathfrak{M}_{a,q}$. Exactly as in I, if α is in $\mathfrak{m}$, there exist a, q such that

$$|\alpha - a/q| \leqq 1/qQ, \qquad P^{1/2} < q \leqq Q - 1. \tag{6}$$

Taking the interval of integration in (4) to be (5), we have

$$r_s(N) = \sum_{q \leqq P^{1/2}} \sum_a \int_{\mathfrak{M}_{a,q}} T^{s_2}(\alpha)U^2(\alpha)e(-N\alpha)\,d\alpha + \int_{\mathfrak{m}} T^{s_2}(\alpha)U^2(\alpha)e(-N\alpha)\,d\alpha. \tag{7}$$

Major arcs.

LEMMA 8.

$$\sum_{q \leqq P^{1/2}} \sum_{a} \int_{\mathfrak{M}_{a,q}} | T^{s_2}(\alpha) - T^{*s_2}(\alpha, a, q) | \; | U(\alpha) |^2 \, d\alpha = O(U^2 P^{s_2 - k - (1/5)}).$$

Proof. If α is in $\mathfrak{M}_{a,q}$, then by Lemmas 5, 6,

$$T^{s_2}(\alpha) - T^{*s_2}(\alpha, a, q) = O(q^{(3/4)+\epsilon} q^{-(s_2-1)/k} \min (P^{s_2-1}, P^{-(s_2-1)(k-1)} \, | \beta |^{-(s_2-1)})).$$

Hence

$$\int_{\mathfrak{M}_{a,q}} | T^{s_2}(\alpha) - T^{*s_2}(\alpha, a, q) | \, d\alpha$$

$$= O(q^{(3/4)+\epsilon-(s_2-1)/k} \int_0^\infty \min(P^{s_2-1}, P^{-(s_2-1)(k-1)} \beta^{-(s_2-1)}) \, d\beta)$$

$$= O(q^{(3/4)+\epsilon-(s_2-1)/k} P^{s_2-1-k}).$$

Thus the sum of the Lemma is

$$O(U^2 P^{s_2-1-k} \sum_{q \leqq P^{1/2}} q q^{(3/4)+\epsilon-(s_2-1)/k}) = O(U^2 P^{s_2-1-k+((11/4)+\epsilon-(s_2-1)/k)/2}).$$

The exponent of P is less than $s_2 - k + (\epsilon/2)$ by an amount

$$1 - (11/8) + (s_2 - 1)/2k = 9/40 \text{ if } k = 5; \quad = 3/8 \text{ if } k = 6.$$

LEMMA 9. *If $\overline{\mathfrak{M}}_{a,q}$ denotes the part of the interval* (5) *which does not belong to $\mathfrak{M}_{a,q}$, then*

$$\sum_{q \leqq P^{1/2}} \sum_{a} \int_{\overline{\mathfrak{M}}_{a,q}} | T^{*}(\alpha, a, q) |^{s_2} \, | U(\alpha) |^2 \, d\alpha = O(U^2).$$

Proof. Since the integrand is a periodic function of α with period 1, we can take the interval of integration to be $q^{-1}Q^{-1} \leqq | \beta | \leqq 1/2$, where $\alpha = a/q + \beta$. By Lemma 5, the sum in question is therefore

$$O(U^2 \sum_{q \leqq P^{1/2}} \sum_{a} \int_{q^{-1}Q^{-1}}^{\infty} q^{-s_2/k} P^{-s_2(k-1)} \beta^{-s_2} d\beta)$$

$$= O(U^2 \sum_{q \leqq P^{1/2}} q q^{-s_2/k} P^{-s_2(k-1)} q^{s_2-1} P^{(s_2-1)(k-1+\delta)})$$

$$= O(U^2 P^{-k+1+(s_2-1)\delta} P^{(1+s_2-(s_2/k))/2}).$$

The exponent of P is

$$-k + 1 + (1 + s_2 - (s_2/k))/2 + (s_2 - 1)\delta = \begin{cases} -(7/10) + 6\delta \text{ if } k = 5; \\ -(1/3) + 9\delta \text{ if } k = 6, \end{cases}$$

and is therefore negative.

Minor arcs.

Lemma 10. *In* $\mathfrak{m}$, $T(\alpha) = O(P^{1-\kappa+\delta})$, *where* $\kappa = 1/2^{k-1}$.

Proof. For any α in $\mathfrak{m}$, there exist a, q satisfying (6).

Case 1. $P^{1/2} < q \leqq P^{1-\delta}$. By Lemma 6,

$$T(\alpha) = O(q^{-1/k}P) = O(P^{1-(1/2k)}) = O(P^{1-k}).$$

Case 2. $P^{1-\delta} < q \leqq Q - 1 < P^{k-1+\delta}$. Lemma 7.

Lemma 11. $\displaystyle\int_{\mathfrak{m}} |T(\alpha)|^{s_2} |U(\alpha)|^2 \, d\alpha = O(U^2 P^{s_2-k-10^{-5}})$.

Proof. Plainly

$$\int |U(\alpha)|^2 \, d\alpha = U.$$

where the integral is taken over any interval of length 1. Also, by Lemma 10, in $\mathfrak{m}$,

$$T^{s_2}(\alpha) = O(P^{s_2(1-\kappa+\delta)}).$$

Hence the integral of the Lemma is

$$O(UP^{s_2(1-\kappa+\delta)}) = O(U^2 P^{s_2(1-\kappa+\delta)-k\mu_{s_1}^{(k)}+\epsilon}),$$

by (3). Using Lemma 2, the exponent of P is

$$s_2 - (s_2/2^{k-1}) - k\mu_{s_1}^{(k)} + s_2\delta + \epsilon < s_2 - k - 10^{-4} + s_2\delta + \epsilon < s_2 - k - 10^{-5}.$$

The singular series.

Lemma 12. $\displaystyle\int_0^1 T^{*s_2}(\alpha, a, q) e(-n\alpha) \, d\alpha = q^{-s_2}(S_{a,q})^{s_2} e_q(-na) R(n)$,

where, for $N - 2P^k < n < N$, $R(n)$ *satisfies*

$$C_1 P^{s_2-k} < R(n) < C_2 P^{s_2-k}. \tag{8}$$

Proof. The integral is

$$q^{-s_2}(S_{a,q})^{s_2} e_q(-na) \int_0^1 I^{s_2}(\beta) e(-n\beta) \, d\beta = q^{-s_2}(S_{a,q})^{s_2} e_q(-na) R(n),$$

where

$$R(n) = k^{-s_2} \sum_{n_1, \ldots, n_{s_2}} (n_1 \cdots n_{s_2})^{-1+(1/k)},$$

and the variables of summation are subject to

$$P^k \leqq n_1, \cdots, n_{s_2} \leqq (2P)^k, \qquad n_1 + \cdots + n_{s_2} = n. \tag{9}$$

Plainly

$$C_3 P^{-(k-1)s_2} R'(n) < R(n) < C_4 P^{-(k-1)s_2} R'(n),$$

where $R'(n)$ denotes the number of solutions of (9). Since $n_1, \cdots, n_{s_2-1}$ determine n_{s_2} uniquely, we have

$$R'(n) < (2P)^{k(s_2-1)}.$$

Also, by (2), $100P^k \leqq N < 101P^k$, so that $98P^k < n < 101P^k$. Hence, if

$$96P^k/(s_2 - 1) \leqq n_1, \cdots, n_{s_2-1} \leqq 97P^k/(s_2 - 1),$$

the value of n_{s_2} determined by the last condition in (9) also satisfies the first condition. Thus

$$R'(n) \geqq (P^k/(s_2 - 1))^{s_2-1}.$$

This establishes Lemma 12.

Let

$$A(n, q) = \sum_a q^{-s_2} (S_{a,q})^{s_2} e_q(-na).$$

Lemma 13. $A(n, q) = O(q^{1-(s_2/k)})$.

Proof. Lemma 3.

Lemma 14. *If* $(q_1, q_2) = 1$, *then* $A(n, q_1 q_2) = A(n, q_1) A(n, q_2)$.

Proof. Landau, Satz 282.

The following notation corresponds to that of Landau, pp. 280-302, except that in some cases we are precluded from using the same symbols.

For any prime p we define γ as follows:

(*a*) if $k = 5$, then $\gamma = 1$ for $p \neq 2, 5$, and $\gamma = 2$ for $p = 2$ or 5;

(*b*) if $k = 6$, then $\gamma = 1$ for $p \neq 2, 3$, and $\gamma = 3$ for $p = 2$, and $\gamma = 2$ for $p = 3$.

For any prime p and any l, n, we denote by $N(p^l, n)$ the number of solutions of

$$x_1^k + \cdots + x_{s_2}^k \equiv n \pmod{p^l}, \qquad 1 \leqq x_i \leqq p^l,$$

in which not all the x's are divisible by p.

Lemma 15. *Let* $k\rho + \sigma$ *be the exact power to which* p *divides* n, *where* $0 \leqq \sigma < k$. *Let*

$$l_0 = \max(\rho k + \sigma + 1, \rho k + \gamma).$$

Then

$$A(n, p^l) = 0 \quad \textit{for} \quad l > l_0,$$

and

$$\chi_p(n) = \sum_{\nu=0}^{\infty} A(n, p^\nu) = p^{(1-s_2)\gamma} N(p^\gamma, 0) \sum_{\tau=0}^{\rho-1} p^{\tau(k-s_2)} + p^{\rho(k-s_2)+\gamma(1-s_2)} N(p^\gamma, n/p^{k\rho}),$$

where, if $\rho = 0$, *the sum on the right is to be read as zero.*

Proof. Landau, Satz 293. (Note that Landau uses P for p^γ, and β instead of ρ).

COROLLARY. *If* $p \nmid n$, *and* $\gamma = 1$, *then* $A(n, p^l) = 0$ *for* $l > 1$.

LEMMA 16. $N(p^\gamma, n) > 0$ *for all* n *and* p.

Proof. By Landau, Satz 301, it suffices to prove that $s_2 \geqq r + 1$, where

$$r = [(p^\gamma - 1)/(p - 1)](k, p - 1).$$

If $k = 5$, we have

$$r = (5, p - 1) \text{ if } p \neq 2,\ 5;\ = 3 \text{ if } p = 2;\ = 6 \text{ if } p = 5;$$

thus $r + 1 \leqq 7 = s_2$. If $k = 6$, we have

$$r = (6, p - 1) \text{ if } p \neq 2, 3;\ = 7 \text{ if } p = 2;\ = 8 \text{ if } p = 3;$$

thus $r + 1 < 10 = s_2$.

LEMMA 17. *For any prime* p *and any* n, $\chi_p(n) > p^{-C_5}$.

Proof. Case 1. Suppose $p^k \nmid n$, so that $\rho = 0$ in Lemma 15. By Lemmas 15, 16,

$$\chi_p(n) = p^{\gamma(1-s_2)} N(p^\gamma, n) \geqq p^{\gamma(1-s_2)} > p^{-C_5}.$$

Case 2. Suppose $p^k | n$, so that $\rho \geqq 1$. By Lemmas 15, 16,

$$\chi_p(n) \geqq p^{(1-s_2)\gamma} N(p^\gamma, 0) \geqq p^{(1-s_2)\gamma} > p^{-C_5}.$$

LEMMA 18. *For any prime* p *and any* n, $| A(n, p) | < C_6 p^{1-(s_2/2)}$.

Proof. Landau, Satz 318.

LEMMA 19. *For any prime* p *and any* n, $\chi_p(n) > 1 - C_7 p^{1-(s_2/2)}$.

Proof. Landau, Satz 322.

LEMMA 20. *For any* n, *the series* $\mathfrak{S}(n) = \sum_{q=1}^{\infty} A(n, q)$ *converges absolutely, and* $\mathfrak{S}(n) > C_8$.

Proof. By the Corollary to Lemma 15, and by Lemma 18, if $p \nmid 30n$,

$$| \chi_p(n) - 1 | = | A(n, p) | < C_6 p^{-5/2}.$$

Hence the product $\prod_p \chi_p(n)$ converges absolutely; therefore, by Lemma 14, so does the series $\sum_{q=1}^{\infty} A(n, q)$, and the two are equal. By Lemmas 17, 19,

$$\prod_p \chi_p(n) > \prod_{p \leqq C_7} p^{-C_5} \prod_{p > C_7} (1 - C_7 p^{-5/2}) = C_8.$$

LEMMA 21. *For* $\eta \geqq 1$, $\sum_{q \geqq \eta} | A(n, q) | = O(\eta^{-1/5} n^{\epsilon})$.

Proof. Case 1. $k = 5$. Any q is representable as $2^{\tau} 5^{\tau'} q_1 q_2$, where

(1) q_1 is *quadratfrei* and not divisible by 2 or 5;

(2) q_2 is composed of prime powers with exponents $\geqq 2$, and is not divisible by 2 or 5;

(3) $(q_1, q_2) = 1$.

By Lemma 13,

$$A(n, 2^{\tau}) = O(2^{-2\tau/5}), \quad A(n, 5^{\tau'}) = O(5^{-2\tau'/5}), \quad A(n, q_2) = O(q_2^{-2/5}).$$

By Lemma 15, if $p^l | q_2$ and $p^{l+1} \nmid q_2$, and $A(n, p^l) \neq 0$, then $l \leqq l_0 = 5\rho + \sigma + 1$, where $5\rho + \sigma$ is the exact power to which p divides n. Since $l \geqq 2$, it follows that $p^l \mid p^{2(l-1)} \mid n^3$. Hence, if $A(n, q_2) \neq 0$, we must have $q_2 \mid n^3$.

By Lemmas 14, 18,

$$| A(n, q_1) | = \prod_{p | q_1} | A(n, p) | < \prod_{p | q_1} C_6 p^{-5/2} = O(q_1^{-(5/2)+\epsilon}).$$

Hence

$$\sum_{q \geqq \eta} | A(n, q) | = O\Bigl(\sum_{q_1} \sum_{\substack{q_2 \\ q_2 | n^3 \\ q_1 q_2 2^{\tau} 5^{\tau'} \geqq \eta}} \sum_{\tau} \sum_{\tau'} q_1^{-(5/2)+\epsilon} q_2^{-2/5} 2^{-2\tau/5} 5^{-2\tau'/5}\Bigr)$$

$$= O\Bigl(\eta^{-1/5} \sum_{q_1=1}^{\infty} \sum_{q_2 | n^3} \sum_{\tau=0}^{\infty} \sum_{\tau'=0}^{\infty} q_1^{-(23/10)+\epsilon} q_2^{-1/5} 2^{-\tau/5} 5^{-\tau'/5}\Bigr)$$

$$= O(\eta^{1/5} n^{\epsilon}).$$

Case 2. $k = 6$. The proof is similar, expressing q as $2^{\tau} 3^{\tau'} q_1 q_2$, and appealing to the same Lemmas. The exponents are more favourable, since s_2/k is now 10/6 instead of 7/5.

Proof of (1). By (7),

$$\begin{aligned} r_s(N) = {} & \sum_{q \leqq P^{1/2}} \sum_{a} \int_0^1 T^{*s_2}(\alpha, a, q) U^2(\alpha) e(-N\alpha)\, d\alpha \\ & - \sum_{q \leqq P^{1/2}} \sum_{a} \int_{\mathfrak{M}_{a,q}} T^{*s_2}(\alpha, a, q) U^2(\alpha) e(-N\alpha)\, d\alpha \\ & + \sum_{q \leqq P^{1/2}} \sum_{a} \int_{\mathfrak{M}_{a,q}} (T^{s_2}(\alpha) - T^{*s_2}(\alpha, a, q)) U^2(\alpha) e(-N\alpha)\, d\alpha \\ & + \int_{\mathfrak{m}} T^{s_2}(\alpha) U^2(\alpha) e(-N\alpha)\, d\alpha. \end{aligned}$$

By Lemma 9, the second sum is $O(U^2)$. By Lemma 8, the third sum is $O(U^2 P^{s_2-k-(1/5)})$. By Lemma 11, the final integral is $O(U^2 P^{s_2-k-10^{-3}})$. Also the first sum, by Lemma 12, is

$$\begin{aligned} & \sum_{k=1}^{U} \sum_{j=1}^{U} \sum_{q \leqq P^{1/2}} \sum_{a} \int_0^1 T^{*s_2}(\alpha, a, q) e(u_k\alpha + u_j\alpha - N\alpha)\, d\alpha \\ & = \sum_{k=1}^{U} \sum_{j=1}^{U} \sum_{q \leqq P^{1/2}} \sum_{a} q^{-s_2} (S_{a,q})^{s_2} e_q(-(N - u_k - u_j)a) R(N - u_k - u_j) \\ & = \sum_{h=1}^{U} \sum_{j=1}^{U} \sum_{q \leqq P^{1/2}} A(N - u_k - u_j, q) R(N - u_k - u_j). \end{aligned}$$

By Lemma 21,

$$\sum_{q \leqq P^{1/2}} A(N - u_k - u_j, q) = \mathfrak{S}(N - u_k - u_j) + O(P^{-(1/10)+\epsilon}).$$

Since $N - u_k - u_j$ lies between $N - 2P^k$ and N, we have

$$C_1 P^{s_2-k} < R(N - u_k - u_j) < C_2 P^{s_2-k},$$

by Lemma 12. Hence

$$r_s(N) = \sum_{h=1}^{U} \sum_{j=1}^{U} \mathfrak{S}(N - u_k - u_j) R(N - u_k - u_j) + O(U^2 P^{s_2-k-10^{-3}}).$$

By Lemma 20, $\mathfrak{S}(N - u_k - u_j) > C_8$. Thus the above sum is greater than

$$C_1 C_8 U^2 P^{s_2-k}.$$

Hence

$$r_s(N) > 0.$$

THE UNIVERSITY,
MANCHESTER, ENGLAND.

A CLASS OF LINEAR INTEGRAL TRANSFORMATIONS.*

By H. L. GARABEDIAN

1. Introduction. In 1924 Silverman [1] studied a class of integral (kernel) transformations which has its analogue in the Hausdorff theory of matrix transformations. In view of certain recent developments in the field of Hausdorff matrix transformations ([2] and [3]) we find it possible now to clarify certain obscurities in Silverman's work. We are also able to make highly significant extensions of Silverman's results on inclusion and equivalence relations among Hausdorff integral transformations.

2. Results of Silverman. Silverman was concerned with the transformation

$$v(x) = \alpha u(x) + \int_0^x k(x, y) u(y) \, dy, \tag{2.1}$$

where $u(x)$ is bounded and integrable, $0 \leqq x \leqq x_1$, where $k(x, y)$ is integrable in y for each positive x, $0 < \delta \leqq y \leqq x$, and where $\int_0^x |k(x, y)| \, dy$ converges for each positive x. The function $k(x, y)$ is called the *kernel* of the transformation (2.1). The transformation is said to be *regular* if the existence of $\lim_{x \to \infty} u(x)$ implies the existence of $\lim_{x \to \infty} v(x)$ and the equality of the limits.

Let $\phi(x)$ be a Hausdorff mass function (*vide infra*) which fulfills the following conditions:

(i) $\phi(x)$ is continuous, $0 \leqq x \leqq 1$,
(ii) $\phi'(x)$ is continuous, $0 < x \leqq 1$,
(iii) $\phi(0) = 0$,
(iv) $\int_0^1 |\phi'(x)| \, dx \leqq M$.

A function $\phi(x)$ satisfying these conditions is called *absolutely regular*. It has been established that the transformation (2.1) with the kernel

$$k(x, y) = \frac{1}{x} \phi'\left(\frac{y}{x}\right), \qquad \alpha = 1 - \phi(1)$$

is a regular transformation when $\phi(x)$ is absolutely regular.

In what follows we shall use the symbol $(H, \phi(x))$ to designate a kernel

---

* Received March 27, 1941.

transformation with the absolutely regular mass function $\phi(x)$, and the symbol $[H, \phi(x)]$ to designate a matrix transformation with the mass function $\phi(x)$ which is regular in the sense of Hausdorff. We shall also use this symbolism in connection with well-known particular transforms. For example, the symbols (C, α) and $[C, \alpha]$ will be used to designate Cesàro summability for integral and matrix transformations respectively.

Silverman's main results involve the integral equations

$$\phi_a(x) = [1 - \phi_c(1)]\phi_b(x) + \phi_c(x) + \int_x^1 \phi_b\left(\frac{x}{s}\right)\phi'_c(s)\,ds, \tag{2.3}$$

$$\phi_a(x) = [1 - \phi_b(1)]\phi_c(x) + \phi_b(x) + \int_x^1 \phi_c\left(\frac{x}{s}\right)\phi'_b(s)\,ds. \tag{2.4}$$

These equations may be expressed symbolically in the form

$$\phi_a(x) = \phi_c(x) \circ \phi_b(x),$$
$$\phi_a(x) = \phi_b(x) \circ \phi_c(x).$$

The following two theorems embody Silverman's principal results.

THEOREM 1. *A sufficient condition in order that* $(H, \phi_a(x)) \supset (H, \phi_b(x))$ *is that there exist an absolutely regular solution* $\phi_c(x)$ *of the equations* (2.3), (2.4).

THEOREM 2. *A sufficient condition that the transformation* $(H, \phi(x))$ *be equivalent to the identical transformation* $v(x) = u(x)$ *is that there exist an absolutely regular solution of the equation* $\phi(x) \circ \omega(x) = 0$.

Silverman used the equation (2.4) to prove that $(H, 2) \approx (C, 2)$, where $(H, 2)$ is the symbol for Hölder integral summability of order 2. He suggested that his method could be used to prove that $(H, k) \approx (C, k)$, $(k = 1, 2, 3, \cdots)$, a result obtained by Landau [4] in 1913; but this technique would be very laborious. An extension of Landau's result to include the case of non-integral orders of $k > 0$ was made by M. Jacob [5] in 1927.

3. Hausdorff matrix transformations. Before discussing Silverman's results we recall certain results in the domain of Hausdorff matrix transformations.

The transformation

$$\sigma_m = \sum_{n=0}^{m} \binom{m}{n} \Delta^{m-n} c_n \cdot s_n,$$

where $\{s_n\}$ is an infinite sequence of numbers and

$$c_n = \int_0^1 x^n d\phi(x),$$

defines a regular Hausdorff method of summation $[H, \phi(x)]$, or a *regular moment sequence* (*regular sequence*), if and only if

$$(3.1)\quad \begin{array}{ll} \text{(i)} & \phi(x) \text{ is of bounded variation on the interval } 0 \leqq x \leqq 1, \\ \text{(ii)} & \phi(x) \text{ is continuous at } x = 0, \text{ and } \phi(0) = 0, \\ \text{(iii)} & \phi(1) = 1, \\ \text{(iv)} & \phi(x) = \tfrac{1}{2}[\phi(x-0) + \phi(x+0)] \text{ if } 0 \leqq x < 1. \end{array}$$

Associated with a regular Hausdorff method $[H, \phi(x)]$ is the function

$$c(z) = \int_0^1 x^z d\phi(x),\ \mathbf{R}(z) \geqq 0,$$

called the *moment function* of the Hausdorff method, and also the function

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots = \int_0^1 \frac{d\phi(u)}{1 - xu},$$

called the *moment generating function* of the Hausdorff method.

We now list together various *equivalent* formulations of the Hausdorff inclusion problem.

THEOREM 3. $[H, \phi_a(x)] \supset [H, \phi_b(x)]$

(i) *if and only if there exists a regular sequence* $\{c_n\}$ *such that*

$$a_n = b_n c_n, \qquad b_n \neq 0, \qquad n = 0, 1, 2, \cdots,$$

(ii) *if and only if there exists a regular moment function* $c(z)$ *such that*

$$a(z) = b(z) c(z), \qquad \mathbf{R}(z) \geqq 0,$$

(iii) *if and only if there exists a regular mass function* $\phi_c(x)$ *such that*

$$f_a(x) = \int_0^1 f_b(vx)\, d\phi_c(v),$$

(iv) *if and only if there exists a regular mass function* $\phi_c(x)$ *such that*

$$\phi_a(x) = \int_0^1 \phi_c\left(\frac{x}{s}\right) d\phi_b(s), \tag{3.1}$$

$$\phi_a(x) = \int_0^1 \phi_b\left(\frac{x}{s}\right) d\phi_c(s), \tag{3.2}$$

for all except at most a countable set of values of x *in the interval* $0 < x < 1$.

A discussion of the nature of the exceptional set in the last formulation may be found in [3].

4. Extensions of Silverman's results. In the discussion of this section we need the following theorem.

THEOREM 4. *Let* $\phi(x)$ *be a bounded monotone non-decreasing function on the interval* $(0,1)$; *then if* $c_n = \int_0^1 x^n d\phi(x)$, $(n = 0, 1, 2, \cdots)$,

$$\lim_{n\to\infty} c_n = \phi(1) - \phi(1-0).$$

This theorem was stated without proof in a recently published paper by the present author and H. S. Wall [6], p. 192. The proof has been included in a paper, by the same authors, as yet unpublished.[1]

This theorem may be extended readily to include the case of a function $\phi(x)$ of limited variation. If $\phi(x)$ is a function of limited variation it is well known that we can write $\phi(x) = \phi_1(x) - \phi_2(x)$, where $\phi_1(x)$ and $\phi_2(x)$ are bounded monotone non-decreasing functions. We then have

$$c_n = \int^1 x^n d\phi(x) = \int_0^1 x^n d\phi_1(x) - \int_0^1 x^n d\phi_2(x),$$

and from Theorem 4 it follows that

$$\begin{aligned}\lim_{n\to\infty} c_n &= \phi_1(1) - \phi_1(1-0) - \phi_2(1) + \phi_2(1-0)\\ &= \phi(1) - \phi(1-0).\end{aligned}$$

We are now in a position to prove the following theorem which is relevant to this investigation.

THEOREM 5. *If* $\{c_n\}$ *is a regular sequence equivalent to or identical with convergence then the associated mass function* $\phi(x)$ *has a discontinuity at* $x = 1$ *equal to* $1 - \phi(1-0)$.

If the sequence $\{c_n\}$ is equivalent to or identical with convergence, then by Theorem 3, formulation (i), the sequence $\{1/c_n\}$ must be a regular sequence. Since all regular sequences are bounded we must then require that $\lim_{n\to\infty} c_n \neq 0$. From the extension of Theorem 4 it now follows that $1 - \phi(1-0) \neq 0$. This completes the proof of our theorem.

In order to establish a more complete connection between the theories of Hausdorff kernel and matrix transformations than is known at the present time, we must revise the conditions of absolute regularity (2.2) in the light of our present knowledge of Hausdorff matrix transformations. From Theorem

---

[1] H. L. Garabedian and H. S. Wall, "Topics in continued fractions and summability," *Northwestern University Studies in Mathematics and the Physical Sciences*, No. 1 (1941), pp. 87-132. Added in proof.

5 we know that mass functions equivalent to or identical with convergence in the domain of Hausdorff matrix transformations are discontinuous at $x = 1$. However, mass functions of this character are not absolutely regular in the sense of Silverman. The conditions of absolute regularity (2.2) may be restated in such a way as to preserve intact the proofs and results of Silverman and to unify the two theories of Hausdorff transformations.

We shall now re-define a function $\phi(x)$ to be absolutely regular provided that

(i) $\phi(x)$ is continuous, $0 \leqq x < 1$, $\lim\limits_{x \to 1^-} \phi(x)$ exists and is equal to $\phi(1-0)$,

(ii) $\phi'(x)$ is continuous, $0 < x < 1$, $\lim\limits_{x \to 1^-} \phi'(x)$ exists and is equal to $\phi'(1-0)$,

(4.1) (iii) $\phi(0) = 0$,

(iv) $\phi(1) = 1$,

(v) $\displaystyle\int_x^1 |\phi'(x)|\, dx \leqq M.$

We now have

$$\alpha = 1 - \phi(1-0),$$

and Silverman's integral equations (2.3) and (2.4) become

$$(4.2) \quad \phi_a(x) = [1 - \phi_c(1-0)]\phi_b(x) + \phi_c(x) + \int_x^1 \phi_b\left(\frac{x}{s}\right)\phi'_c(s)\, ds,$$

$$(4.3) \quad \phi_a(x) = [1 - \phi_b(1-0)]\phi_c(x) + \phi_b(x) + \int_x^1 \phi_c\left(\frac{x}{s}\right)\phi'_b(s)\, ds.$$

We observe that if $\phi(x)$ is continuous at $x = 1$ we have $\phi(1) = \phi(1-0)$ and $\alpha = 0$. We also observe that if $\phi(x)$ fulfills the conditions of absolute regularity (4.1) then $\phi(x)$ satisfies the conditions of regularity (3.1) and is hence regular in the sense of Hausdorff.

To find the connecting link between the Hausdorff theories of matrix and integral transformations it remains only to investigate briefly the integral equations (3.1) and (3.2).

First of all we write these equations in the form

$$(4.4) \qquad \phi_a(x) = \phi_b(x) + \int_x^1 \phi_c\left(\frac{x}{t}\right) d\phi_b(t),$$

$$(4.5) \qquad \phi_a(x) = \phi_c(x) + \int_x^1 \phi_b\left(\frac{x}{t}\right) d\phi_c(t).$$

Effecting an integration by parts in each of these equations we obtain

$$(4.6)\quad \phi_a(x) = [1-\phi_c(1-0)]\phi_b(x) + \phi_c(x) + x\int_x^1 \phi_b(t)\phi'_c\left(\frac{x}{t}\right)\frac{1}{t^2}\,dt,$$

$$(4.7)\quad \phi_a(x) = [1-\phi_b(1-0)]\phi_c(x) + \phi_b(x) + x\int_x^1 \phi_c(t)\phi'_b\left(\frac{x}{t}\right)\frac{1}{t^2}\,dt.$$

If in each of the integrals in (4.6) and (4.7) we make the change of variable defined by the transformation $ts = x$ and then invert the order of integration, the equations (4.6) and (4.7) become (4.2) and (4.3) respectively. Thus, the theories of Hausdorff integral and matrix transformations are coextensive for mass functions which are absolutely regular.

The implications of the preceding discussion are highly significant. For example, all of the inclusion and equivalence relationships which are known among Hausdorff matrix transformations associated with absolutely regular mass functions are now valid in the domain of Hausdorff integral transformations. Moreover, we have now available all four of the formulations of the Hausdorff inclusion problem listed in Theorem 3 to use in investigating inclusion and equivalence relationships between kernel transformations associated with absolutely regular mass functions. In this connection we observe that ordinarily it is easier to use formulations (i) and (ii) than formulation (iv) of Theorem 3 in investigating inclusion and equivalence relationships among Hausdorff transformations.

In order to exhibit some of the hitherto unknown inclusion and equivalence relations between Hausdorff integral transformations we list several mass functions which satisfy the conditions of absolute regularity (4.1):

$$\phi_1(x) = \frac{1}{\Gamma(\alpha)}\int_0^x \left(\log\frac{1}{t}\right)^{\alpha-1} dt, \qquad \alpha > 0,$$

$$\phi_2(x) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)}\int_0^x t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt, \qquad \gamma > \alpha > 0,$$

$$\phi_3(x) = 1 - e^{x/(x-1)}.$$

The function $\phi_1(x)$ defines Hölder summability $[H, \alpha]$. The function $\phi_2(x)$ defines hypergeometric summability $[\mathrm{H}, \alpha, 1, \gamma]$ [6]. In particular, summability $[H, 1, 1, \alpha+1]$ is identical with Cesàro summability $[C, \alpha]$. The function $\phi_3(x)$ defines a method of summation which includes all of the Cesàro methods [7].

That $(C, \alpha) \approx (H, \alpha)$, $\alpha > 0$, was proved by M. Jacob [5] in 1927, as previously mentioned, using rather tedious methods. This result also follows from the well-known equivalence $[C, \alpha] \approx [H, \alpha]$, $\alpha > 0$. A number of inclusion and equivalence relations are known among the hypergeometric matrix transformations [6]. All of these are now valid in the domain of kernel transformations. We list a few of these:

$$(H, \alpha, 1, \gamma_1) \subset (H, \alpha, 1, \gamma_2), \gamma_2 > \gamma_1 > \alpha > 0,$$
$$(H, \alpha_2, 1, \gamma) \subset (H, \alpha_1, 1, \gamma), \gamma > \alpha_2 > \alpha_1 > 0,$$
$$(H, \alpha, 1, \alpha + \beta) \approx (C, \beta), \alpha, \beta > 0.$$

The last of these relations has been recently proved [8] in the domain of matrix transformations. Finally, we have the interesting result

$$(H, \phi_8(x)) \supset (C, \alpha), \alpha > 0.$$

NORTHWESTERN UNIVERSITY.

REFERENCES

1. L. L. Silverman, "The equivalence of certain regular transformations," *Transactions of the American Mathematical Society*, vol. 26 (1924), pp. 101-112.

2. E. Hille and J. D. Tamarkin, "Questions of relative inclusion in the domain of Hausdorff means," *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 573-577.

3. H. L. Garabedian, E. Hille, and H. S. Wall, "Formulations of the Hausdorff inclusion problem," *Duke Mathematical Journal*, vol. 8 (1941), pp. 193-213.

4. E. Landau, "Die Identität des Cesàroschen und Hölderschen Grenzwertes für Integrale," *Sächsische Berichte*, vol. 65 (1913), pp. 131-138.

5. M. Jacob, "Über die Äquivalenz der Cesàroschen und der Hölderschen Mittel für Integrale bei gleicher reeller Ordnung $k > 0$," *Mathematische Zeitschrift*, vol. 27 (1927), pp. 672-682.

6. H. L. Garabedian and H. S. Wall, "Hausdorff methods of summation and continued fractions," *Transactions of the American Mathematical Society*, vol. 48 (1940), pp. 185-207.

7. H. L. Garabedian, "Hausdorff methods of summation which include all of the Cesàro methods," accepted for publication by the *Bulletin of the American Mathematical Society*.

8. H. L. Garabedian, "Relations between hypergeometric methods of summation," presented to the American Mathematical Society, December 1940; cf. *Bulletin of the American Mathematical Society*, vol. 47 (1941), abstract number 24, p. 36.

EXTENSION AND REDUCTION THEOREMS FOR CERTAIN TYPES OF CONTINUOUS TRANSFORMATIONS.*

By D. D. MILLER.

---

1. Introduction. In recent years several types of single-valued continuous transformations have been defined and investigated by various authors. Of these transformations, some of the most interesting are the interior transformations defined by S. Stoïlow [1][1] the non-alternating transformations defined by G. T. Whyburn [2], the non-separating transformations of J. F. Wardwell [3], and the monotone transformations studied by Whyburn [2] and others. Stronger and weaker forms of the non-alternating and non-separating transformations have been considered by E. P. Vance [4]. These transformations will be defined as they are used in later sections of this paper.

In this paper we study what might be called "hereditary" properties of such transformations. That is, if P is one of the properties mentioned above (interiority, monotoneity, etc.), we inquire under what circumstances a transformation having property P on a space A will have property P on every subcontinuum of A, or on certain special subcontinua (arc-sets and cyclic elements) of A. The conditions for this "inheritance" of a property by a subcontinuum or set of subcontinua may take the form of conditions on the transformations or of conditions on the subcontinua. In addition to these reduction theorems, we give an extension theorem valid for several types of transformation.

The author wishes to acknowledge his indebtedness to Professor W. L. Ayres, who suggested the investigation of this topic, for his advice and encouragement during the preparation of the paper.

Unless otherwise specified, A will denote, throughout this paper, a Peano space (= continuous image of an interval = compact locally connected continuum), and T a single-valued continuous transformation. If a transformation $T(A) = B$ has property P, and if R is a subcontinuum of A, T will be said to *have property P on R* if $T(R) = S$, where S is a subset of R, has property P. Definitions of cyclic elements, arc-sets, etc. may be found in a paper by Kuratowski and Whyburn [5]; we assume that the reader is familiar with the definitions and results of this paper.

---

* Received March 18, 1941; presented to the American Mathematical Society, April 12, 1940.

[1] Figures in square brackets refer to the bibliography at the end of the paper.

2. Transformations possessing a property hereditarily. In this section we obtain necessary and sufficient conditions on transformations of certain types on a space A in order that these transformations shall possess their defining properties on every subcontinuum of A.

Our first theorem concerns monotone transformations, *i. e.*, transformations under which the inverse image of each point of the transform is connected.

THEOREM 2. 1. *A necessary and sufficient condition that a transformation $T(A) = B$ be monotone on every subcontinuum of A is that the inverse image of each point of B be an arc-set.*[2]

This theorem follows readily from results of Whyburn [6, pp. 311-312].

A transformation is said to be *non-alternating* on A if, given any two points of the transform, neither inverse image separates the other in A. (Throughout this paper we use "separate" in the strong sense; *i. e.*, a set X is said to separate a set M in A if $A - X$ is the sum of two disjoint sets, each closed relative to $A - X$, and each containing at least one point of M.)

THEOREM 2. 2. *A transformation $T(A) = B$ is monotone on every subcontinuum of A if and only if it is non-alternating on every subcontinuum of A.*

Since every monotone transformation is non-alternating, the necessity is immediate.

To prove the sufficiency, we use the characterization of transformations monotone on every subcontinuum of A given by the preceding theorem. Let y be a point of B, and suppose that $T^{-1}(y)$ is not an arc-set. Then there is an arc α in A with only its endpoints in $T^{-1}(y)$. But T is obviously not non-alternating on α, contrary to the hypothesis that T is non-alternating on every subcontinuum of A. Hence every $T^{-1}(y)$ must be an arc-set, and T must, therefore, be monotone on every subcontinuum of A.

We now obtain a necessary and sufficient condition, of different character from the condition of Theorem 2. 1, in order that a transformation be monotone (or, by the preceding theorem, non-alternating) on every subcontinuum of the space. We shall say that a transformation is *constant* on a set E if it carries E into a single point.

THEOREM 2. 3. *A necessary and sufficient condition that a monotone transformation $T(A) = B$ be monotone on every subcontinuum of A is that T be homeomorphic or constant on each cyclic element of A.*

---

[2] We use the term *arc-set* in the sense of *closed arc-set* (= *A-set* in the terminology of [6]).

The condition is necessary. Obviously T is constant on each degenerate cyclic element ($=$ a single point). Now let C be a true cyclic element of A. If T is not homeomorphic on C, then, for some point y of $T(C)$, $T^{-1}(y) \cdot C$ is non-degenerate; and if T is not constant on C, there exists a point x of $T(C)$ distinct from y. Let p and r be distinct points of $T^{-1}(y) \cdot C$, and let q be a point of $T^{-1}(x) \cdot C$. By Ayres' Three-Point Theorem [7, p. 130, Th. 3], there exists an arc pqr in C. But, contrary to the hypothesis that T is monotone, and hence non-alternating, on every subcontinuum of A, T then fails to be non-alternating on the subcontinuum pqr, for $T^{-1}(y)$ intersects two components of $pqr - T^{-1}(x)$, one in p and the other in r.

The condition is sufficient.[3] Let T be monotone on A and homeomorphic or constant on each cyclic element of A. Suppose that there is a continuum K in A and a point y of B such that $K \cdot T^{-1}(y) = M + N$, where the sets M and N are disjoint and closed. There is an arc α in A, with endpoints m and n, such that $\alpha \cdot M = m$ and $\alpha \cdot N = n$. Since K and $T^{-1}(y)$ are continua, it is easy to see that no point separates m from n in A, whence $m + n$ lies in a cyclic element E. Since T is thus not homeomorphic on E, we must have $E \subset T^{-1}(y)$. Now EK is connected, but $EK = E \cdot K \cdot T^{-1}(y) = EM + EN$, and the summands are mutually separated. This is a contradiction.

For the necessity part of Theorem 2. 3, the condition that T be monotone on A is evidently superfluous. However, without this condition the sufficiency part of the theorem is not true. For example, let A be an arc xy and let T be a transformation which identifies x and y but is homeomorphic elsewhere on A.

As an immediate consequence of Theorems 2. 2 and 2. 3 we have

THEOREM 2. 4. *A necessary and sufficient condition that a non-alternating transformation $T(A) = B$ be non-alternating on every subcontinuum of A is that T be homeomorphic or constant on each cyclic element of A.*

COROLLARY. *A necessary and sufficient condition that every non-alternating transformation on a Peano continuum A be non-alternating on every subcontinuum of A is that A be a dendrite.*

The sufficiency of the condition follows at once from Theorem 2. 4, for every single-valued transformation on a dendrite is necessarily constant on every cyclic element. Conversely, on a non-acyclic Peano continuum we can always define a non-alternating transformation which, on some cyclic element, is neither homeomorphic nor constant, and which in consequence, by the above theorem, must fail to be non-alternating on some subcontinuum. Such a transformation may be defined as one which carries into a point an arc in some true cyclic element of the continuum, and which is a homeomorphism elsewhere; this transformation is obviously monotone, and hence non-alternating.

---

[3] The author is indebted to the referee for this abbreviated proof of the sufficiency.

A transformation $T(A) = B$ is said to be *weakly non-alternating* provided, for any two distinct points x and y of B, $T^{-1}(x)$ separates $T^{-1}(y)$ in A only if some single point of $T^{-1}(x)$ separates $T^{-1}(y)$ in A. The transformation would be called *completely non-alternating* if, for any two points x and y (not necessarily distinct) of B, and any closed subset K of $T^{-1}(x)$, K did not separate $T^{-1}(y)$ in A.

The necessity part of Theorem 2.4 fails if the hypothesis that T be non-alternating on every subcontinuum of A is replaced by the weaker hypothesis that T be weakly non-alternating on every subcontinuum of A. For example, let A be a circle and let T identify two points on the circle (T being a homeomorphism elsewhere on A); T is then weakly non-alternating on every subcontinuum of A (and, in fact, completely non-alternating on A), but is neither homeomorphic nor constant on the cyclic element A.

It will follow as a trivial corollary of Theorem 2.5 that the sufficiency part of Theorem 2.4 holds if "weakly non-alternating" replaces "non-alternating" throughout the statement of the theorem. However, if only the first occurrence of "non-alternating" is replaced by "weakly non-alternating," the theorem is no longer true. For example, consider the transformation which folds a line segment about its midpoint; this transformation is weakly non-alternating on the segment and is constant on every cyclic element, but is not non-alternating even on the whole segment. Since every transformation completely non-alternating on every subcontinuum of a Peano space is obviously a homeomorphism, it is apparent that replacement of the second occurrence of "non-alternating" by "completely non-alternating" will render the sufficiency theorem false unless the first occurrence of "non-alternating" is similarly replaced, in which case the theorem is trivially true.

A transformation $T(A) = B$ is called *non-separating* if the inverse image of no point of B separates A. It is *weakly non-separating* if an inverse image can separate A only if one of its points separates A. It is *completely non-separating* if no closed subset of any inverse image separates A.

Results similar to the above for the various types of non-separating transformations, insofar as they are not contained in the foregoing remarks, are to be found in Theorems 2.5, 2.6, and 2.7. There is one exception, however: it is impossible to define a transformation completely non-separating on every subcontinuum of a Peano space.

As a preliminary to the proof of Theorem 2.5 we prove the following

LEMMA. *If H is a subcontinuum of a Peano space A, and K is a closed set which separates H between two points p and q, then there exists a cyclic element C of A such that KC separates H between p and q.*

Consider the cyclic chain $C(p,q)$ from p to q. Since every cut point of $C(p,q)$ is a cut point of A, and hence a cyclic element, if K contains a cut point of $C(p,q)$ the lemma is proved. Let us suppose, then, that $K \cdot Z = 0$, where Z is the set of all cut points of $C(p,q)$. Since $H \cdot C(p,q)$ is a subcontinuum of $C(p,q)$ containing $p+q$, $Z \subset H \cdot C(p,q)$. But, by hypothesis, $H \cdot C(p,q) - K = H_p + H_q$, where H_p and H_q are mutually separated sets containing p and q, respectively. Hence $Z = Z_p + Z_q$, where Z_p and Z_q are mutually separated and contain p and q, respectively. Since $Z + p + q$ is closed, Z_p and Z_q are closed. The cyclic elements in $C(p,q)$ being linearly ordered between p and q, let p' be the last point of Z_p in the order from p to q, and let q' be the first point of Z_q in the order from p' to q. Since $p' \neq q'$, p' and q' must be the limit points in a certain cyclic element C of the components of $C(p,q) - C$ containing p and q, respectively. The points p' and q' being separated in H by K, they must then be separated by KC, whence p and q are separated in H by KC.

THEOREM 2.5. *If $T(A) = B$ is homeomorphic or constant on each cyclic element of A, then T is weakly non-separating on every subcontinuum of A.*

Suppose that there exists a subcontinuum H of A on which T is not weakly non-separating; *i. e.*, suppose that there exists a point x of $T(H)$ such that $T^{-1}(x)$ separates H while no single point of $T^{-1}(x)$ separates H. By the foregoing lemma, there exists a cyclic element C of A such that $C \cdot T^{-1}(x)$ separates H. Now T cannot be homeomorphic on C, for if it were then $C \cdot T^{-1}(x)$ would be a single point, and we have supposed that no single point of $T^{-1}(x)$ separates H. On the other hand, T cannot be constant on C; for in this case C would separate H, and, C being a cyclic element, C would contain a cut point of A which would separate H, contrary to supposition. But, by hypothesis, T is either homeomorphic or constant on C. Hence, by contradiction, T is weakly non-separating on every subcontinuum of A.

The converse of this theorem is false; for if the space A be taken to be a circle, any transformation carrying a semicircle into a single point, and homeomorphic elsewhere on A, is weakly non-separating on every subcontinuum of A but is neither homeomorphic nor constant on the cyclic element A. However, the hypothesis that T is weakly non-separating on every subcontinuum of A does yield the conclusion embodied in

THEOREM 2.6. *If $T(A) = B$ is weakly non-separating on every subcontinuum of A, then T is monotone on each cyclic element of A (in fact, $T^{-1}(x) \cdot E$ is arcwise connected for every point x of B and every cyclic element E of A).*

Let x be any point of B and E any cyclic element of A such that $T^{-1}(x) \cdot E \neq 0$. If $T^{-1}(x) \cdot E$ is a single point, our conclusion follows trivially. Hence we suppose $T^{-1}(x) \cdot E$ to be non-degenerate and let p and q be distinct points of $T^{-1}(x) \cdot E$. Since $p + q$ lies in a true cyclic element E, there exists a simple closed curve J in E containing $p + q$; and J is the sum of two arcs, α_1 and α_2, having only $p + q$ in common. Since T is, by hypothesis, weakly non-separating on J, and since no single point can separate a simple closed curve, it follows that $J - T^{-1}(x)$ is connected. Hence $T^{-1}(x)$ contains one of the arcs α_i from p to q, whence $T^{-1}(x) \cdot E$ is arcwise connected.

That the converse of this theorem is false is easily seen by considering a θ-curve A (*i. e.*, two points, p and q, and three arcs which join p to q and each pair of which have in common only $p + q$) and a transformation which carries one of the arcs pq into a single point and is homeomorphic elsewhere on A. Each inverse image is arcwise connected, but the transformation is not weakly non-separating even on A.

THEOREM 2.7. *A necessary and sufficient condition that a non-separating transformation $T(A) = B$ be non-separating on every subcontinuum of A is that B be a single point.*

The sufficiency is obvious. To prove the necessity of the condition, suppose that there exist distinct points x and y of B. Let p be a point of $T^{-1}(x)$ and q a point of $T^{-1}(y)$. Since A is a Peano space, there exists an arc pq in A; and since $T^{-1}(x)$ and $T^{-1}(y)$ are disjoint and closed, $pq - T^{-1}(x) - T^{-1}(y) \neq 0$. Let t be a point of $pq - T^{-1}(x) - T^{-1}(y)$; then t is an interior point, and hence a cut point, of pq. But t is a point of $T^{-1}(z)$ for some point z of B; and since $t \cdot T^{-1}(x) = t \cdot T^{-1}(y) = 0$ we have $x \neq z \neq y$. Hence $p \cdot T^{-1}(z) = q \cdot T^{-1}(z) = 0$, so that $T^{-1}(z)$ cuts pq, a subcontinuum of A, contrary to hypothesis.

THEOREM 2.8. *A necessary and sufficient condition that an interior transformation $T(A) = B$ be interior on every subcontinuum of A is that T be homeomorphic or constant on A.*

The sufficiency of the condition is obvious. To prove the necessity, suppose that B is not a single point but that there exist points x_1 and x_2 of A such that $T(x_1) = T(x_2)$. Since A is a Peano space, there exists at least one arc in A joining x_1 and x_2; let α be such an arc. We consider separately the mutually exclusive cases in which (1) there exists a point u of α such that $T(x_1u) \neq T(x_2u)$, where x_iu denotes the subarc of α joining x_i and u $(i = 1, 2)$, and (2) $T(x_1u) = T(x_2u)$, for all points u of α.

Case 1. In this case, since $T(x_1u) \neq T(x_2u)$, there exists a point s of x_1u such that $T(s)$ is not a point of $T(x_2u)$ (or else there exists a point of x_2u whose transform does not belong to $T(x_1u)$; let us suppose the former). Now let p_1 be the first point on α, in the order from s to x_1, such that $T(p_1)$ is a point of $T(x_2u)$ (there exists such a point, for $T^{-1}T(x_2u)$ is closed and at least x_1 itself satisfies the condition). Let p_2 be any point of x_2u such that $T(p_2) = T(p_1)$, let p_1s be the subarc of α from p_1 to s, let $U = T(p_1s - s)$, and let A_1 be the subarc p_1ux_2 of α. Then $p_1s - s$ is a set open relative to A_1. Let $T(A_1) = B_1$. We proceed to show that $T(p_1s - s)$ is not open in B_1, whence the hypothesis that T is interior on every subcontinuum of A is contradicted.

Since $T(p_1) = T(p_2)$, $T(p_1) \subset U \cdot T(x_2u)$; and since p_1 is the only point common to $p_1s - s$ and $T^{-1}T(x_2u)$, $U \cdot T(x_2u) \subset T(p_1)$. Hence $U \cdot T(x_2u) = T(p_1)$. Therefore, since $x_2u \subset A_1$, $T(x_2u) - T(p_1) \subset B_1 - U$. But p_2 is a limit point of a subset M of x_2u, whence $T(p_2)$ is, by the continuity of T, a limit point of $T(M)$, and $T(M) \subset T(x_2u) - T(p_1) \subset B_1 - U$. Therefore, since $T(p_1) = T(p_2)$, the point $T(p_1)$ of U is a limit point of $B_1 - U$, whence U is not open relative to B_1.

Case 2. In this case we suppose that $T(x_1u) = T(x_2u)$ for every point u of the arc α joining x_1 and x_2. If $T(\alpha)$ is a single point, T is obviously interior on α. But we are supposing that B is not a single point, so that there exists a point y in A such that $T(y) \neq T(\alpha)$. Let β be an arc joining α to y. Then T is not interior on the subcontinuum $\alpha + \beta$, contrary to hypothesis.

On the other hand, if we suppose that $T(\alpha)$ is not a single point, *i. e.*, if we suppose there exists a point v of α such that $T(v) \neq T(x_1) = T(x_2)$, then $\rho(T(v), T(x_1)) = \epsilon$, where $\rho(T(v), T(x_1))$ is the distance between $T(v)$ and $T(x_1)$ and ϵ is a positive number. Since T is a continuous transformation, there exists a point u of α so close to x_1 that diam $T(x_1u) < \epsilon$, whence $T(x_1u)$ cannot contain $T(v)$. Hence v is a point of x_2u and $T(x_1u) \neq T(x_2u)$, contrary to supposition.

3. Conditions on the subcontinua. In this section we obtain necessary and sufficient conditions on a subcontinuum R of a Peano space A in order that every transformation having one of the aforementioned properties on A shall have the same property on R. In order to state the theorems in the strongest possible form, we shall in some cases state separately the necessary and the sufficient conditions.

THEOREM 3.1. *In order that every monotone transformation $T(A) = B$ be monotone on a given subcontinuum R of A, it is necessary and sufficient that R be an arc-set.*

The condition is necessary. If R is not an arc-set, then, for some pair of points x and y in R, there exists an arc xy not lying wholly in R. Let T be a transformation carrying xy into a single point p and homeomorphic on $A - xy$. Then T is monotone on A but not on R; for an arc xy is irreducibly connected between x and y, whence $xy \cdot R (= T^{-1}(p) \cdot R)$ is not connected.

The condition is sufficient. Since T is monotone on A, $T^{-1}(x)$ is connected for every point x of B. But, since R is an arc-set, RC is connected for every connected set C in A. Hence $T^{-1}(x) \cdot R$ is connected for every point x of B, and T is therefore monotone on R.

In view of the obvious fact that a Peano space is a dendrite if and only if all its subcontinua are arc-sets, we have at once the

COROLLARY. *A necessary and sufficient condition that every monotone transformation on A be monotone on every subcontinuum of A is that A be a dendrite.*

THEOREM 3.2. *A non-alternating transformation $T(A) = B$ is non-alternating on every arc-set in A.*

Let R be any arc-set in A, and suppose that T is not non-alternating on R. Then there exist points x and y of $T(R)$ such that $T^{-1}(x)$ separates R between some two points p and q of $T^{-1}(y)$. But since R is an arc-set, any set which separates two points in R must separate these points in A. Hence $T^{-1}(x)$ must separate p from q in A, which contradicts the hypothesis that T is non-alternating on A.

The following converse of this theorem is also true: if every non-alternating transformation $T(A) = B$ is non-alternating on a subcontinuum R of A, then R is an arc-set. In this statement, however, we can weaken the hypothesis by replacing the first occurrence of "non-alternating" by "completely non-alternating," without destroying the validity of the conclusion. From this follows, by virtue of Theorem 3.2, the equivalence of the apparently stronger and weaker forms of the hypothesis of the converse. Before proceeding to the proof of this converse, we establish two preliminary lemmas.

LEMMA A. *If a closed set M irreducibly separates a Peano space A between two points p and q, no proper subset of M separates M in A.*

Let C_p and C_q be the components of $A - M$ containing p and q respectively. Since M separates A irreducibly between p and q, M is the common boundary of C_p and C_q [8, p. 48, Th. 7], whence $M \subset \overline{C_p}$. Suppose that there exists a proper subset M' of M which separates M in A; *i. e.*, suppose that there exist mutually separated sets M'_1 and M'_2 such that $A - M' = M'_1 + M'_2 \neq 0$ and and $M \cdot M'_1 \neq 0 \neq M \cdot M'_2$. Since $C_p \subset A - M \subset A - M'$, and since C_p is connected, C_p is contained either in M'_1 or in M'_2, say in M'_1. Hence $\overline{C_p} \subset \overline{M'_1}$, whence $M \subset \overline{M'_1}$. But this implies, since M'_1 and M'_2 are mutually separated, that $M \cdot M'_2 = 0$, contrary to our supposition.

LEMMA B. *If C is a continuum without cut points such that for some two points p and q of C there exist three mutually exclusive connected subsets of C, K_1, K_2, and K_3, each having both p and q as limit points, and if p' and q' are points lying in different components of $C - p - q$, then $C - p' - q'$ is connected.*

Since C has no cut points, every component of $C - p' - q'$ must have both p' and q' as limit points. Hence, if C_i is a component of $C - p' - q'$ containing neither p nor q, $C_i + p' + q'$ is a connected set joining p' and q' in $C - p - q$. But this is impossible, for by hypothesis p' and q' lie in different components of $C - p - q$. Hence every component of $C - p' - q'$ contains either p or q. But $K_i \cdot (p' + q') = 0$ for at least one of the K_i $(i = 1, 2, 3)$, and the connected set $K_i + p + q$ thus joins p and q in $C - p' - q'$, whence p and q lie in the same component of $C - p' - q'$. Hence $C - p' - q'$ has only one component, and thus is connected.

THEOREM 3. 3. *If every completely non-alternating transformation $T(A) = B$ is non-alternating on a subcontinuum R of A, then R is an arc-set.*

Suppose R is not an arc-set. Then there exists an arc α, with end points p_1 and p_2, such that $R \cdot \alpha = p_1 + p_2$. Thus p_1 and p_2 are joined by the continua α and R, having in common only $p_1 + p_2$, whence α lies in some true cyclic element C of A. As will be pointed out later, we shall lose no generality by assuming at this point that the space A is cyclicly connected, so that $C = A$.

Let K be a closed set which irreducibly separates [8, Th. 8] A between p_1 and p_2. Since K and α are closed sets, $K \cdot \alpha$ is closed, whence, R being closed, $\rho(K \cdot \alpha, R) = d > 0$. Hence there exists an open set U (*viz.*, the ϵ-neighborhood of $K \cdot \alpha$, where $\epsilon < d$) such that $U \supset K \cdot \alpha$ and $UR = 0$. Let $K' = K - U$.

Consider now the components C_i of $A - p_1 - p_2$ such that C_i contains a component C_i^* of $R - p_1 - p_2$ such that C_i^* has both p_1 and p_2 as limit points. Since A is locally connected, there is only a finite number[4] of such components C_i. We now consider separately Case 1, in which there is only one such component (that there is at least one is obvious), and Case 2, in which there is more than one.

Case 1. We define the transformation T as follows: $T(C_1 \cdot K) = z$, $T(p_1 + p_2) = t$, where z and t are distinct points, and T is homeomorphic elsewhere on A. Now T is not non-alternating on R, for C_1K' separates p_1 from p_2 in R (since $UR = 0$), whence $T^{-1}(z)$ separates $T^{-1}(t)$ in R. But T is completely non-alternating on A, for (1) no closed subset of C_1K' separates p_1 from p_2 in A (since U contains $K \cdot \alpha$, whence $K' \cdot \alpha = 0$), (2) no closed subset of C_1K' separates C_1K' in A (for any such separating subset of C_1K' would separate K in A, which is impossible by Lemma A), (3) no closed subset of $p_1 + p_2$ separates C_1K' in A (since $C_1 \supset C_1K'$), (4) no closed subset of $p_1 + p_2$ separates $p_1 + p_2$, and (5) no single point can separate either $p_1 + p_2$ or C_1K' (since A was assumed to be cyclic).

Case 2. If there is more than one of the components C_i, each such component must intersect K', whence $p_1 + p_2$ separates $K' \cdot \Sigma C_i$ in A, and a transformation defined as in Case 1 fails even to be non-alternating on A. To avoid this difficulty we redefine our transformation T as follows.

Let $K''_i = K' \cdot C_i$ and $K'' = \Sigma K''_i = K' \cdot \Sigma C_i$, noting that this is a finite sum. Then each K''_i is closed, whence K'' is closed, and therefore $\rho(p_1 + p_2, K'') = d > 0$. The space A being locally connected, there exists a positive number δ such that every point in $U_\delta(p_i)$[5] can be joined to p_i by an arc lying within $U_d(p_i)$ $(i = 1, 2)$. Let p'_i be any point of $U_\delta(p_i) \cdot C_i$, and let β_i be an arc $p_ip'_i$ lying within $U_d(p_i)$ $(i = 1, 2)$. Now K'' separates p'_1 from p'_2 in R, for K'' separates p_1 from p_2 in R and $K'' \cdot U_d(p_i) = 0$ $(i = 1, 2)$. Hence if we define T as the transformation which carries K'' into a point z and $p'_1 + p'_2$ into another point t, and is homeomorphic elsewhere on A, T is not non-alternating on R. But T is completely non-alternating on A, for (1) no closed subset of K'' separates p'_1 from p'_2 in A (since the connected set $\beta_1 + \alpha + \beta_2$ does not intersect K''), (2) no closed subset of K'' separates K'' in A (for the reason cited in (2) of Case 1), (3) no closed subset of $p'_1 + p'_2$ separates K'' in A (in fact, if in Lemma B we let $p = p_1$, $q = p_2$, $p' = p'_1$, $q' = p'_2$, $K_1 = C_1$, $K_2 = C_2$, and $K_3 = \alpha - p_1 - p_2$, we see that $p'_1 + p'_2$

---

[4] For if the number of such components C_i is infinite, let $\{x_i\}$ be a sequence of points such that x_i is a point of C_i and $p_1 \neq x_i \neq p_2$ $(i = 1, 2, \cdots)$. Then, since A is compact, the sequence has a limit point x, and in the neighborhood of x the space fails to be locally connected.

[5] The symbol $U_r(x)$ denotes the neighborhood of x of radius r.

does not cut A at all), (4) no closed subset of $p'_1 + p'_2$ separates $p'_1 + p'_2$, and (5) no single point separates $p'_1 + p'_2$ or K'' or itself in A (since A is assumed to be cyclic).

If the space were not assumed to be cyclicly connected, the foregoing steps would remain unaltered if T were defined to be a homeomorphism on $A - C$.

We have thus defined, on the supposition that R is not an arc-set, a transformation completely non-alternating on A but not non-alternating on R, contrary to hypothesis. Hence R is an arc-set.

THEOREM 3.4. *In order that every non-separating transformation $T(A) = B$ be non-separating on a given subcontinuum R of A, it is necessary and sufficient that R be an arc-set.*

The sufficiency is proved in the same way as Theorem 3.2. To prove the necessity, we shall make slight modifications in the proof of Theorem 3.3. Using the notation of that proof, and supposing as before that R is not an arc-set, consider the set $A - K'' = C_{p_1p_2} + \Sigma C'_i$, where $C_{p_1p_2}$ is the component of $A - K''$ containing p_1 and p_2 (these points lie in the same component of $A - K''$ because they are joined by the arc α, which does not intersect K'') and the C'_i are the other components of $A - K''$. Let $K''' = K'' + \Sigma C'_i$. Now all the limit points in $A - C'_i$ of any component C'_i must lie in K'', whence[6] $F(C'_i) \subset K''$ for all C'_i, and therefore $\Sigma F(C'_i) \subset K''$. But K'' is closed, so that $\overline{\Sigma F(C'_i)} \subset K''$, and, since the space is Peanian, $F(\Sigma C'_i) \subset \overline{\Sigma F(C'_i)} \subset K''$ [9, p. 155]. Therefore K''' is closed, and $A - K'''$ $(= C_{p_1p_2})$ is connected, whence K''' does not separate A. On the other hand, ΣC_i contains neither p_1 nor p_2 and K'' separates R between p_1 and p_2 (as we have seen in the proof of Theorem 3.3), so that K''' separates R between p_1 and p_2. Therefore, assuming A to be cyclic, if T be defined to carry K''' into a point z and to be homeomorphic elsewhere on A, T is non-separating on A but not non-separating on R, contrary to hypothesis. If A is not cyclic and C is the true cyclic element of A containing $p_1 + p_2$, let T carry any component of $A - C$ into its limit point in C; the result then follows as before. The supposition that R is not an arc-set thus leads to a contradiction and the theorem is proved.

Whether the theorem remains true if the first occurrence of "non-separating" be replaced by "completely non-separating," is not yet known.

Theorems 3.1–3.4 may be summarized in the

COROLLARY. *If A is a Peano space and R is a subcontinuum of A, the*

---

[6] The symbol $F(x)$ denotes the *boundary* of the set X, *i. e.*, the set $\overline{X} \cdot \overline{A - X}$.

following properties are equivalent: (1) *R is an arc-set,* (2) *every transformation monotone on A is monotone on R,* (3) *every transformation non-alternating on A is non-alternating on R,* (4) *every transformation completely non-alternating on A is non-alternating on R,* (5) *every transformation non-separating on A is non-separating on R.*

In the same way that Theorem 3.2 was established, we can prove

THEOREM 3.5. *If a transformation $T(A) = B$ is completely non-alternating* (resp. *completely non-separating, weakly non-alternating, weakly non-separating*), *it has the same property on every arc-set in A.*

Some of the converses of Theorem 3.4 are also true. Since Theorem 3.3 holds, it is true *a fortiori* that if, in that theorem, "non-alternating" or "weakly non-alternating" be substituted for "completely non-alternating," or if "completely non-alternating" be substituted for "non-alternating," or both, the theorem still holds. However, if "weakly non-alternating" be substituted for "non-alternating" in Theorem 3.3, then even if "weakly non-alternating" be substituted for "completely non-alternating" the resulting theorem is false. As indicated in the remark following Theorem 3.4, we do not know whether that theorem remains true if "completely non-separating" be substituted for "non-separating." In other respects the foregoing remarks concerning the various non-alternating transformations apply without change to the corresponding types of non-separating transformations.

4. Extension theorems. In this section we establish an extension theorem for monotone, non-alternating, and non-separating transformations; *i. e.*, a theorem which asserts that a transformation having one of these properties on certain subcontinua of the space has the same property on the whole space.

W. L. Ayres has defined [10, p. 568] a *basic set* of a Peano space A to be a subset of A such that each point of A lies on an arc joining two points of the subset. In this section we show that transformations which are monotone, non-alternating, or non-separating on all the cyclic chains joining pairs of points of a basic set possess their defining properties on the whole space.

THEOREM 4.1. *If K is a basic set of A, and if $T(A) = B$ is a transformation monotone on $C(x, y)$ for every pair of points x, y in K, then T is monotone on A.*

Suppose that T is not monotone on A; then there exists a point b of B such that $T^{-1}(b) = U + V$, where U and V are non-vacuous mutually separated sets. Let u and v be points of U and V, respectively; then no connected subset of $T^{-1}(b)$ contains $u+v$. We consider separately the cases in which (1) $u+v$ lies in some true cyclic element of A, and (2) u and v lie in distinct cyclic elements of A.

Case 1. If u and v lie in some true cyclic element C, let s be a non-cut point of A contained in C. Since K is a basic set of A, there exist points x and y in K such that $s \in C(x, y)$. But if an arc-set contains a non-cut point of a cyclic element, it contains the cyclic element.[7] Hence $C(x, y) \supset C \supset u+v$. But then $T^{-1}(b)$ is not connected in $C(x, y)$ and T is not monotone on $C(x, y)$, contrary to hypothesis.

Case 2. If u and v do not lie in the same cyclic element, there exist cyclic elements C_u and C_v, containing u and v, respectively, and such that either $C_u \cdot C_v = 0$ or $C_u \cdot C_v$ is a single point. We proceed to find points x and y of K such that $C(x, y) \supset u+v$. If C_u is a node (*i. e.*, contains exactly one cut point of A), then [11, Th. 8] C_u contains a point of K which is not a cut point of A; call this point x. On the other hand, if C_u is not a node (*i. e.*, contains more than one cut point of A), let c be the particular cut point in C_u which is the limit point of the component of $A - C_u$ containing C_v, and take as x any point of K lying in a component [10, Th. 11] M of $A - C_u$ whose limit point in C_u is a point d distinct from c.

Now we may select a point y in exactly the way that x was chosen, merely interchanging u and v in the above paragraph. Thus $C(x, y)$ contains either a non-cut point of C_i or two cut points of C_i $(i = u, v)$, and in either case $C(x, y) \supset C_i$. Hence $C(x, y) \supset C_u + C_v \supset u + v$, whence T is not monotone on $C(x, y)$, contrary to hypothesis.

We now prove, in much simpler fashion, that the same theorem holds for non-alternating and non-separating transformations.

THEOREM 4. 2. *If K is a basic set of A, and if $T(A) = B$ is a transformation non-alternating (non-separating) on $C(x, y)$ for every pair of points x, y in K, then T is non-alternating* (resp. *non-separating*) *on A.*

If T is not non-alternating on A, there exist points p and q of B such that $T^{-1}(p)$ separates A between some two points u and v of $T^{-1}(q)$, whence $T^{-1}(p)$ separates u from v in any connected subset of A containing $u+v$. But we have seen in the proof of Theorem 4. 1 that there exist points x and y

---

[7] Every cyclic chain is an arc-set, and it follows immediately from the definition of the latter that a non-degenerate arc-set contained in a cyclic element C must contain C; and an arc-set which is not contained in C but contains a non-cut point in C must also contain a cut point in C, and hence, containing two points of C, must contain C.

of K such that $C(x, y) \supset u + v$. Hence $T^{-1}(p)$ separates $C(x, y)$ between u and v, contrary to the hypothesis that T is non-alternating on $C(x, y)$ for all points x, y in K. In the same way, T may be shown to be non-separating on A if it is non-separating on each $C(x, y)$.

OHIO UNIVERSITY,
ATHENS, OHIO.

BIBLIOGRAPHY

1. S. Stoïlow, "Sur les transformations continues et la topologie des fonctions analytiques," *Ann. Scient. de l'Ecole Norm. Sup.*, vol. 45 (1928), pp. 347-382.

2. G. T. Whyburn, "Non-alternating transformations," *American Journal of Mathematics*, vol. 56 (1934), pp. 294-302.

3. J. F. Wardwell, "Non-separating transformations," *Duke Mathematical Journal*, vol. 2 (1936), pp. 745-750.

4. E. P. Vance, "Generalizations of non-alternating and non-separating transformations," *Duke Mathematical Journal*, vol. 6 (1940), pp. 66-79.

5. C. Kuratowski and G. T. Whyburn, "Sur les éléments cycliques et leurs applications," *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331.

6. G. T. Whyburn, "Arc-preserving transformations," *American Journal of Mathematics*, vol. 58 (1936), pp. 305-312.

7. W. L. Ayres, "Continuous curves which are cyclicly connected," *Bull. de l'Acad. Polonaise des Sc. et des Lettres, Classe des Sc. Math. et Nat., Série A: Sc. Math.* (1928), pp. 127-142.

8. G. T. Whyburn, "Concerning irreducible cuttings of continua," *Fundamenta Mathematicae*, vol. 13 (1929), pp. 42-57. Cf. the work of Kuratowski cited therein.

9. F. Hausdorff, *Mengenlehre*, 3rd ed. (Berlin: 1935).

10. W. L. Ayres, "Concerning the arc-curves and basic sets of a continuous curve," *Transactions of the American Mathematical Society*, vol. 30 (1928), pp. 567-578.

11. W. L. Ayres, "Concerning the arc-curves and basic sets of a continuous curve, second paper," *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 595-612.

CALCUTTA UNIV CENTRAL LIBRARY

THE PROBLEM OF PLATEAU IN HYPERBOLIC SPACE.[*]

CALCUTTA UNIVERSITY CENTRAL LIBRARY

By A. T. Lonseth.

1. Preliminary considerations. The spaces of constant Riemannian curvature[1] fall into three types according as the curvature is zero, negative, or positive. In the first (Euclidean) case, minimal surfaces and the Plateau problem have been studied intensively.[2] Here the second (hyperbolic) case is considered. The central result obtained is a proof of the fact that if a Jordan curve in three dimensions is capable of bounding a surface of finite area in the sense of hyperbolic geometry it actually bounds a surface which is minimal in the sense of hyperbolic geometry.

The three-dimensional hyperbolic geometry whose curvature $K = -1$ can be represented on the half-space of (x, y, z) in which $z > 0$. Its distance element is then given by

$$ds^2 = (dx^2 + dy^2 + dz^2)z^{-2} \tag{1}$$

("Poincaré's metric"). Geodesics are circles orthogonal to the (x, y)-plane.

A vector function $\mathfrak{x}(u, v) : x(u, v),\ y(u, v),\ z(u, v)$ will define a *surface* in $z > 0$ if the three coördinate functions are continuous on the closed unit circle of a (u, v)-plane; $z(u, v) > 0$, so that it possesses a positive lower bound. We shall say that $\mathfrak{x}(u, v)$ is an *admissible* vector if the first partial derivatives of its components are continuous in the open unit circle, and are summable with their squares thereon.

If the expressions E, F and G are defined as the coefficients of the first fundamental form of $\mathfrak{x}(u, v)$:

---

[*] Received September 9, 1939; in revised form, July 20, 1941; presented to the American Mathematical Society, April 15, 1939.

This paper is a revision, without any essential modifications, of the author's University of California thesis (May, 1939). The author is deeply indebted to Professor Hans Lewy for his interest and assistance.

Several valuable improvements were suggested by the referee. The most noteworthy of these is incorporated in the bounding procedure of § 4, which replaces a much less elegant method.

[1] Bianchi-Lukat: *Vorlesungen über Differential-Geometrie*, Leipzig, 1899; pp. 575 and 584.

[2] For work done up to 1933, see Radó: *On the Problem of Plateau*, Berlin, 1933. For more recent investigations, see Courant, "The existence of minimal surfaces of given topological structure under prescribed boundary conditions," *Acta Mathematica*, vol. 72 (1940), pp. 51-98.

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2$$

then, from (1),

$$(2)\qquad \begin{aligned} E &= \mathfrak{x}_u^2 z^{-2} = (x_u^2 + y_u^2 + z_u^2) z^{-2}, \\ F &= \mathfrak{x}_u \mathfrak{x}_v z^{-2} = (x_u x_v + y_u y_v + z_u z_v) z^{-2} \\ G &= \mathfrak{x}_v^2 z^{-2} = (x_v^2 + y_v^2 + z_v^2) z^{-2}. \end{aligned}$$

For an admissible vector $\mathfrak{x}(u, v)$, the *hyperbolic Dirichlet integral*

$$\begin{aligned} B[\mathfrak{x}] &= \iint_{u^2+v^2<1} (E + G)\,du\,dv \\ &= \iint_{u^2+v^2<1} (x_u^2 + x_v^2 + y_u^2 + y_v^2 + z_u^2 + z_v^2) z^{-2}\,du\,dv \end{aligned}$$

exists. Noting that

$$\sqrt{EG - F^2} \leqq \sqrt{EG} \leqq \tfrac{1}{2}(E + G),$$

we conclude that the *hyperbolic area integral*

$$A[\mathfrak{x}] = \iint_{u^2+v^2<1} \sqrt{EG - F^2}\,du\,dv$$

exists and satisfies the inequality

$$(3)\qquad A[\mathfrak{x}] \leqq \tfrac{1}{2} B[\mathfrak{x}].$$

In (3), the equality occurs if and only if $E \equiv G$, $F \equiv 0$.

In terms of spherical coördinate functions $r(u, v)$, $\theta(u, v)$, $\phi(u, v)$ obtained from $x(u, v)$, $y(u, v)$, $z(u, v)$ by the transformation

$$(4)\qquad x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r \cos\theta,$$

we find that

$$(5)\qquad B[\mathfrak{x}] = B_1[\mathfrak{x}] + B_2[\mathfrak{x}],$$

where

$$B_1[\mathfrak{x}] = \iint_{u^2+v^2<1} (r_u^2 + r_v^2) r^{-2} \sec^2\theta\,du\,dv,$$

$$B_2[\mathfrak{x}] = \iint_{u^2+v^2<1} \{(\theta_u^2 + \theta_v^2) \sec^2\theta + (\phi_u^2 + \phi_v^2 \tan^2\theta)\}\,du\,dv.$$

The fact that $B_2[\mathfrak{x}]$ is independent of $r(u, v)$ will be of use later.

Certain notations will be employed rather frequently. A function will be said to be of class $C^{(n)}$ in a domain D if all its derivatives of order n exist and are continuous in D. The open circle whose center is at P and whose radius is R will be denoted by $C(P, R)$; the closed disk by $\bar{C}(P, R)$; and the circumference by $C^*(P,R)$. Thus in particular the open unit circle in the (u,v)-plane is $C(O,1)$.

2. Statement of the problem. If a surface is defined by a vector of class C'' in $C(O,1)$, whose coördinate functions satisfy in $C(O,1)$ the Euler equations of the problem $A[\mathfrak{x}] = \text{minimum}$, we shall call it a *hyperbolic minimal surface.*

A particular class of hyperbolic "minimal varieties" consists of the spheres (including planes) which are orthogonal to the fundamental plane $z = 0$. These are totally goedesic (i. e., their geodesics are also geodesics of the enveloping space), so their minimal character follows from a general theorem.[3]

Let there be given in (x,y,z) a closed, continuous curve Γ without double points (Jordan curve). We shall say that a continuous vector $\mathfrak{x}(u,v)$ *spans* Γ if the equations $x = x(u,v)$, $y = y(u,v)$, $z = z(u,v)$ map $C^*(O,1)$ onto Γ. The boundary mapping, though continuous, need not be $(1-1)$.

We can now state the *problem of Plateau for hyperbolic space*: given a Jordan curve Γ in hyperbolic space, it is required to find a hyperbolic minimal surface whose coördinate vector $\mathfrak{x}(u,v)$ *spans* Γ. This problem will be solved for any curve Γ which can be spanned by an admissible vector, as defined in § 1.

In view of the minimal character of totally geodesic surfaces, a solution of this Plateau problem would be obvious in case Γ should lie on a sphere or plane orthogonal to $z = 0$: namely, that portion of the sphere (or plane) "interior" to Γ. This remark is the key to a bounding procedure employed in § 4.

We proceed to the analytical details in the general case. Writing $z = \exp \zeta$,

$$(2') \qquad \begin{aligned} E &= \exp(-2\zeta)(x_u{}^2 + y_u{}^2) + \zeta_u{}^2, \\ F &= \exp(-2\zeta)(x_u x_v + y_u y_v) + \zeta_u \zeta_v, \\ G &= \exp(-2\zeta)(x_v{}^2 + y_v{}^2) + \zeta_v{}^2 \end{aligned}$$

The Euler equations $(\mathcal{E}^*)$ for the problem $A[\mathfrak{x}] = \text{minimum}$, obtained by independent variations of x, y and ζ are:[4]

$$(\mathcal{E}^*) \qquad \begin{aligned} &\frac{\partial}{\partial u}\left[\tfrac{1}{2}(GE_{x_u} - 2FF_{x_u})W^{-1}\right] + \frac{\partial}{\partial v}\left[\tfrac{1}{2}(EG_{x_v} - 2FF_{x_v})W^{-1}\right] = 0, \\ &\frac{\partial}{\partial u}\left[\tfrac{1}{2}(GE_{y_u} - 2FF_{y_u})W^{-1}\right] + \frac{\partial}{\partial v}\left[\tfrac{1}{2}(EG_{y_v} - 2FF_{y_v})W^{-1}\right] = 0, \\ &\frac{\partial}{\partial u}\left[\tfrac{1}{2}(GE_{\zeta_u} - 2FF_{\zeta_u})W^{-1}\right] + \frac{\partial}{\partial v}\left[\tfrac{1}{2}(EG_{\zeta_v} - 2FF_{\zeta_v})W^{-1}\right] \\ &\qquad\qquad - \tfrac{1}{2}(EG_\zeta + GE_\zeta - 2FF_\zeta)W^{-1} = 0, \end{aligned}$$

---

[3] L. P. Eisenhart, *Riemannian Geometry*, Princeton, 1926, p. 184.

[4] O. Bolza, *Vorlesungen über Variationsrechnung*, Leipzig and Berlin, 1909, pp. 655, 667.

where $W = (EG - F^2)^{\frac{1}{2}}$. If we suppose $E - G \equiv F \equiv 0$ in $C(O,1)$—the case of isothermic parameters—equations $(\mathcal{E}^*)$ become, since then $W \equiv E \equiv G$, on removing the factor $\frac{1}{2}$,

$$\frac{\partial}{\partial u}[E x_u] + \frac{\partial}{\partial v}[G x_v] = 0,$$

$$\frac{\partial}{\partial u}[E y_u] + \frac{\partial}{\partial v}[G y_v] = 0,$$

$$\frac{\partial}{\partial u}[E \zeta_u] + \frac{\partial}{\partial u}[G \zeta_v] - [E + G] = 0.$$

By $(2')$, these finally reduce to

$$(\mathcal{E}) \qquad \begin{aligned} &\Delta x - 2(\zeta_u x_u + \zeta_v x_v) = 0, \\ &\Delta y - 2(\zeta_u y_u + \zeta_v y_v) = 0, \\ &\Delta \zeta + \exp(-2\zeta)(x_u^2 + x_v^2 + y_u^2 + y_v^2) = 0, \end{aligned}$$

where Δ is the Laplacian operator

$$\Delta = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}.$$

For future reference these equations $(\mathcal{E})$ will be denoted by $\mathcal{E}(x;\zeta)$, $\mathcal{E}(y;\zeta)$ and $\mathcal{E}(\zeta;x,y)$ respectively.

It is easily seen that if $E - G \equiv F \equiv 0$ and if equations $(\mathcal{E})$ hold, then equations $(\mathcal{E}^*)$ also hold. Accordingly, the condition on a hyperbolic minimal surface, that its coördinate vector satisfy the Euler equations $(\mathcal{E}^*)$ for the problem $A[\mathfrak{x}] =$ minimum, can be replaced by the following (apparently more stringent) pair of conditions:

1. $E - G \equiv F \equiv 0$ *in* $C(O,1)$;
2. *equations* $(\mathcal{E})$ *hold in* $C(O,1)$.

The problem of Plateau becomes:[5] given a Jordan curve Γ for which $z > 0$, it is required to find a vector $\mathfrak{x}(u,v)$: $x(u,v)$, $y(u,v)$, $z(u,v)$, where $z(u,v) > 0$, such that

1. $\mathfrak{x}(u,v)$ *is continuous on* $\bar{C}(O,1)$ *and is of class* C'' *in* $C(O,1)$;
2. $\mathfrak{x}(u,v)$ *maps* $C^*(O,1)$ *continuously onto* Γ;
3. *the quantities* E, F *and* G *defined by* (2) *satisfy* $E - G \equiv F \equiv 0$ *in* $C(O,1)$;
4. *the equations* $(\mathcal{E})$, *in which* $\zeta = \log z$, *are satisfied in* $C(O,1)$.

---

[5] Cf. Radó, *l. c.*[2], pp. 31, 32.

3. The Plateau problem as a minimum problem for $B[\mathfrak{x}]$. The system of equations $(\mathcal{E})$ is precisely the system of Euler equations arising from the problem $B[\mathfrak{x}] = \text{minimum}$. Hence if $\mathfrak{x}(u, v)$

1. *spans* Γ *and is continuous on* $\bar{C}(O, 1)$;
2. *is of class* C'' *in* $C(O, 1)$;
3. *is such that* $B[\mathfrak{x}] = \text{g.l.b.}\ \{B[\mathfrak{x}]\}$, *where* $\{\mathfrak{x}\}$ *is the class of admissible vectors spanning* Γ; then $\mathfrak{x}(u, v)$ must satisfy $(\mathcal{E})$ in $C(O, 1)$. Concerning Γ we now make explicit the further assumption that it can be spanned by an admissible vector, so that $\{\mathfrak{x}\}$ is not empty and $\{B[\mathfrak{x}]\}$ consequently possesses a finite greatest lower bound.

Further, if $\mathfrak{x}(u, v)$ has the three properties just listed, it follows necessarily, that $E - G \equiv F \equiv 0$ in $C(O, 1)$. Thus if an $\mathfrak{x}(u, v)$ can be found with these properties, it defines a hyperbolic minimal surface; for $E - G \equiv F \equiv 0$ and equations $(\mathcal{E})$ hold, so that the variational equations $(\mathcal{E}^*)$ are satisfied in $C(O, 1)$. The problem of finding a vector function satisfying conditions 1, 2, and 3 we shall call *the minimum problem for* $B[\mathfrak{x}]$. A solution of this problem is a solution of the problem of Plateau in hyperbolic space.

The proof of the isothermic relation $E - G \equiv F \equiv 0$ under hypotheses 1, 2, 3 on $\mathfrak{x}(u, v)$ does not differ essentially from proofs of the corresponding theorem in Euclidean space, as given by Radó[6] and by Courant.[7] Radó's proof requires only the obvious replacement of the ordinary by the hyperbolic Dirichlet integral; Courant's, in addition, use of equations $(\mathcal{E})$ instead of Laplace's equation to establish analyticity of the function of a complex variable

$$\phi(w) = \phi(u + iv) = E - G - 2iF$$

within $C(O, 1)$. The vanishing of $\phi(w)$ identically in $C(O, 1)$ then follows from the hypothesis that $\mathfrak{x}(u, v)$ minimizes the B-integral among variations of the mapping of $C''(O, 1)$ onto Γ. We omit the details.

In what follows, the minimum problem for $B[\mathfrak{x}]$ is solved. The method of attack is fundamentally the *direct method* of the calculus of variations: a minimal sequence of admissible vectors spanning Γ is improved until finally a vector with properties 1, 2, 3 is obtained as a limit. Major improvements are effected by the *indirect method*, involving solutions of the variational equations for analytic approximating boundary values. The difficulties encountered are due to the non-linearity of equations ($\mathcal{E}$), which complicates improvement and convergence of minimal sequences, and even makes a modified type of minimal sequence seem necessary.

---

[6] T. Radó, *l. c.*[2], p. 97, Lemma.

[7] R. Courant, "Plateau's problem and Dirichlet's principle," ***Annals of Mathematics***, vol. 38 (1937), pp. 679-724; p. 693.

Use will be made of the Euler equation obtained from $B[\mathfrak{x}] =$ minimum by variation of $r(u, v)$, leaving $\theta(u, v)$ and $\phi(u, v)$ fixed. Writing

$$\lambda = \log r, \tag{6}$$

we have from (5)

$$B_1[\mathfrak{x}] = \int\int (\lambda_u^2 + \lambda_v^2) \sec^2 \theta \, du \, dv;$$

as has been remarked, $B_2[\mathfrak{x}]$ does not involve $r(u, v)$. The Euler equation for $\lambda(u, v)$ is then

$$\Delta\lambda - 2 \tan \theta (\theta_u \lambda_u + \theta_v \lambda_v) = 0,$$

which we write as $\mathcal{E}(\lambda; \theta)$.

4. A bounded, analytic minimal sequence. Consider the totality $\{\mathfrak{x}\}$ of admissible vectors spanning a given Jordan curve Γ in hyperbolic space. Since $B[\mathfrak{x}] \geqq 0$, the set of numbers $\{B[\mathfrak{x}]\}$ possesses a non-negative greatest lower bound, finite because of our assumption that $\{\mathfrak{x}\}$ is not empty. Write

$$b = \text{g. l. b. } \{B[\mathfrak{x}]\}. \tag{7}$$

We can pick from $\{\mathfrak{x}\}$ an infinite sequence $\{\mathfrak{x}_n\}$ such that $\lim B[\mathfrak{x}_n] = b$.

Instead of proceeding to improve this *minimal sequence,* we construct a *minimal sequence in approximation to* Γ. Suppose that $\epsilon > 0$, and that M is a continuous mapping of $C^*(O, 1)$ onto Γ such that there exists an admissible vector with boundary values M. Consider the totality $\{\mathfrak{x} : M, \epsilon\}$ of vectors whose values on $C^*(O, 1)$ differ from M by less than ϵ. Write

$$\beta(M, \epsilon) = \text{g. l. b.} \{B[\mathfrak{x} : M, \epsilon]\}.$$

Now evidently $\beta(M, \epsilon)$ does not decrease as ϵ decreases. Further, $\beta(M, \epsilon) \leqq B[\mathfrak{x}_m]$, where $\mathfrak{x}_m$ is some admissible vector whose values on $C^*(O, 1)$ are precisely those of M. Consequently $\beta(M, \epsilon)$ has a limit as $\epsilon \rightarrow 0$: $\lim \beta(M, \epsilon) = \beta(M)$.

Now consider the totality $\{M\}$ of mappings which carry $C^*(O, 1)$ continuously into Γ. The set of non-negative numbers $\{\beta(M)\}$ has a finite, non-negative greatest lower bound:

$$\beta = \text{g. l. b.} \{\beta(M)\}. \tag{8}$$

From their definitions, $\beta \leqq b$. It will follow from the existence proof that $\beta = b$.

By (8), a sequence $\{M_n\}$ of mappings can be found such that $\lim \beta(M_n) = \beta$. If $\{\epsilon_n\}$ is a null sequence of positive numbers, we can pick a sequence $\{\mathfrak{x}_n\}$ of vectors such that

1. *on* $C^*(O,1)$, $\mathfrak{x}_n(u,v)$ *differs from* M_n *by less than* ϵ_n;
2. $\lim B[\mathfrak{x}_n] = \beta$.

Such a sequence $\{\mathfrak{x}_n\}$ we shall call a *minimal sequence for* $B[\mathfrak{x}]$ *in approximation* to Γ, or briefly, *minimal sequence.* The term *minimal sequence* will from now on be used in this sense only.

Nothing is known about $\{\mathfrak{x}_n\}$ except "admissibility" of its elements and the two limit properties above. Our first improvement consists in obtaining from this primitive $\{\mathfrak{x}_n\}$ a minimal sequence of uniformly bounded vectors, with $\{z_n(u,v)\}$, moreover, bounded uniformly away from zero. Geometrically, all surfaces defined by the new minimal sequence lie above a plane parallel to the (x,y)-plane, and within a sphere whose center is at the origin.

The continuous, closed curve Γ lies in the half-space $z > 0$. It is consequently completely above a plane $z = m$ and within a sphere $x^2 + y^2 + z^2 < \mathfrak{R}^2$, where $\mathfrak{R} > m > 0$. The boundary values of a minimal sequence in approximation are not necessarily so restricted; but if we construct the sequence with $\epsilon_n = m/8n$, we shall have on $C^*(O,1)$ the inequalities, for all n:

$$z_n > 7m/8, \tag{9}$$

$$x_n^2 + y_n^2 + z_n^2 < (\mathfrak{R} + m/8)^2. \tag{10}$$

From such an $\{\mathfrak{x}_n\}$ we shall now construct a new sequence $\{\mathfrak{x}_n\}$ with the properties:

1. $\mathfrak{x}_n$ *is an admissible vector;*
2. $\{\mathfrak{x}_n\}$ *is a minimal sequence, with* $\epsilon_n = m/8n$;

for all n, *and for* (u,v) *in* $\bar{C}(O,1)$,

3. $z_n(u,v) \geqq 3m/4$;
4. $[\mathfrak{x}_n(u,v)]^2 \leqq (\mathfrak{R} + m/4)^2$.

Two lemmas[8] on certain deformations of $\mathfrak{x}_n$ will be used.

LEMMA 4.1. *Let* $g(t)$ *be a function of class* C' *for* $0 < t < \infty$ *such that*

---

[8] For these, the author has to thank the referee. The method of bounding originally used consisted in cutting off the vectors above by a sphere and below by a plane, and then smoothing by averaging over small circles; the radii of these circles, fixed for each n, tend to zero with $1/n$ in a suitable way.

(i) $g(t) = t$ *for* $0 < t < \log(\mathfrak{R} + m/8)$;

(ii) $g(t) \leqq \log(\mathfrak{R} + m/4)$;

(iii) $0 \leqq g'(t) \leqq 1$.

Suppose $\mathfrak{x}$ *expressed in terms of coördinates* λ, θ, ϕ, (*cf.* (5) *and* (6)). *If* $\mathfrak{x}_n$ *belongs to a minimal sequence with* $\epsilon_n = m/8n$, *the new vector* $\bar{\mathfrak{x}}_n$: $\bar{\lambda}_n$, $\bar{\theta}_n$, $\bar{\phi}_n$, *where* $\bar{\lambda}_n = g(\lambda_n)$, *coincides with* $\mathfrak{x}_n$ *on* $C^*(O,1)$. *On* $\bar{C}(O,1)$, $\bar{\mathfrak{x}}_n{}^2 \leqq (\mathfrak{R} + m/8)^2$. *Further,* $\bar{\mathfrak{x}}_n$ *is admissible, and* $B[\bar{\mathfrak{x}}_n] \leqq B[\mathfrak{x}_n]$.

Such a $g(t)$ is readily constructed. By (i) and (10), boundary values are unchanged. The spherical upper bound follows from (ii). Continuity of $g'(t)$ assures admissibility of $\bar{\mathfrak{x}}_n$.

The only part of $B[\mathfrak{x}_n]$ affected is $B_1[\mathfrak{x}_n]$. Dropping subscripts n for the moment, and integrating over $C(O,1)$, we obtain

$$\begin{aligned} B_1[\bar{\mathfrak{x}}] &= \int\!\!\int (\bar{\lambda}_u{}^2 + \bar{\lambda}_v{}^2)\sec^2\theta\, du dv \\ &= \int\!\!\int [g'(\lambda)]^2(\lambda_u{}^2 + \lambda_v{}^2)\sec^2\theta\, du\, dv \\ &\leqq B_1[\mathfrak{x}], \end{aligned}$$

by (iii).

LEMMA 4.2. *Let* $h(t)$ *be a function of class* C' *for* $0 < t < \infty$ *such that*

(i) $h(t) = t$ *for* $t \geqq 7m/8$;

(ii) $h(t) \geqq 3m/4$;

(iii) $0 \leqq h'(t) \leqq 1$.

If $\bar{\mathfrak{x}}_n$: $\bar{x}_n$, $\bar{y}_n$, $\bar{z}_n$ *is the spherically bounded vector of* LEMMA 4.1, *the new vector* $\mathfrak{x}_n$: $x^*{}_n = \bar{x}_n$, $y^*{}_n = \bar{y}_n$, $z^*{}_n = h(\bar{z}_n)$ *coincides with* $\bar{\mathfrak{x}}_n$ (*and* $\mathfrak{x}_n$) *on* $C^*(O,1)$. *On* $\bar{C}(O,1)$, $z_n \geqq 3m/4$ *and* $\mathfrak{x}^*{}_n{}^2 \leqq (\mathfrak{R} + m)^2$. *Further,* $\mathfrak{x}^*{}_n$ *is admissible, and* $B[\mathfrak{x}^*{}_n] \leqq B[\bar{\mathfrak{x}}_n] \leqq B[\mathfrak{x}_n]$.

Like $g(t)$ in Lemma 4.1, $h(t)$ is easily constructed.

By (i) and (9), boundary values are preserved. By (ii), $z_n \geqq 3m/4$. The spherical bound is seen from the fact that no point on the surface is moved up more than $3m/4$. Continuity of $h'(t)$ assures admissibility of $\mathfrak{x}^*{}_n$. Dropping subscripts as before,

$$\begin{aligned} B[\mathfrak{x}^*] &= \int\!\!\int \{x_u{}^2 + x_v{}^2 + y_u{}^2 + y_v{}^2 + h'(z)^2(z_u{}^2 + z_v{}^2)\}z^{-2}\, dudv \\ &\leqq B[\mathfrak{x}], \end{aligned}$$

We can evidently get the spherical upper bound $(\mathfrak{R} + m/4)^2$ by taking a larger $\mathfrak{R}$.

We shall now replace this bounded $\{\mathfrak{x}^*_n\}$ by an analytic minimal sequence which also is bounded above and below.

The vector $\mathfrak{x}^*_n$ is continuous on $\bar{C}(O, 1)$. We can, consequently, pick a sequence of positive numbers $k_n < 1$ such that, defining $\mathfrak{x}_n(u, v) = \mathfrak{x}^*_n(k_n u, k_n v)$,

1. $\lim k_n = 1$;
2. on $C^*(O, 1)$, $| \mathfrak{x}^*_n - \mathfrak{x}_n | < \epsilon_n = m/8n$ in each coördinate;
3. $| B[\mathfrak{x}^*_n] - B[\mathfrak{x}_n] | < \epsilon_n$.

This new $\{\mathfrak{x}_n\}$ is a minimal sequence, differing from the original by less than $m/8n$ on $C^*(O, 1)$. The vector $\mathfrak{x}_n$ is of class C' on the circle $C(O, k_n^{-1})$, which includes $\bar{C}(O, 1)$. Therefore on $\bar{C}(O, \tfrac{1}{2}(1 + k_n^{-1}))$ we can approximate uniformly to $\mathfrak{x}_n$, $\mathfrak{x}_{nu}$ and $\mathfrak{x}_{nv}$ by a polynomial vector (again called $\mathfrak{x}_n$), the error being less than ϵ_n in each coördinate, both for the components and for their first derivatives.[9] It is easily verified, using the uniform boundedness of $\{B[\mathfrak{x}^*_n]\}$, that $\{\mathfrak{x}_n\}$ is a minimal sequence. On $C^*(O, 1)$ there is approximation within $m/2n$ to the original boundary mapping M_n. Everywhere on $\bar{C}(O, 1)$, and for all n,

$$z_n > m/2; \tag{11}$$

$$x_n^2 + y_n^2 + z_n^2 < (\mathcal{R} + m/2)^2 < (\mathcal{R} + m)^2 \tag{12}$$

We are now in position to apply a selection theorem of Courant's,[10] which will give us a minimal sequence of analytic, bounded vectors whose boundary values converge uniformly to a mapping of $C^*(O, 1)$ onto Γ. First we note that $B[\mathfrak{x}]$ is invariant under conformal mapping; this evidently goes through precisely as does the invariance of the ordinary Dirichlet integral. Accordingly we can construct, by conformal mappings of $\bar{C}(O, 1)$ into itself, a new minimal sequence satisfying a "three-point condition": the mapping M_n, to which $\mathfrak{x}_n$ approximates, carries three distinct fixed points on $C^*(O, 1)$ into three distinct fixed points on Γ, independently of n. For all vectors $\mathfrak{x}_n$, the ordinary Dirichlet integral

$$D[\mathfrak{x}] = \iint\limits_{C(O,1)} (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\, du\, dv$$

is uniformly bounded, since

$$D[\mathfrak{x}_n] < (m/2)^{-2} B[\mathfrak{x}_n] < (m/2)^{-2} M,$$

---

[9] Courant-Hilbert, *Methoden der Mathematischen Physik*, Bd. I, Berlin, 1931, p. 57. Hereafter this book will be referred to as Courant-Hilbert I, and the second volume (Berlin, 1937) as Courant-Hilbert II.

[10] R. Courant, *l. c.*[7], p. 692.

M being the uniform upper bound for $B[\mathfrak{x}_n]$. Hence Courant's theorem can be applied: the boundary values of $\{\mathfrak{x}_n\}$ are equicontinuous, so we may pick a subsequence (again $\{\mathfrak{x}_n\}$) converging uniformly on $C^*(O,1)$. The limit mapping is continuous, and is monotone in the sense of Radó. We shall eventually show that it is also reciprocally $(1-1)$.

5. On solutions of the variational equations $(\mathcal{E})$. Further improvements will be gained by replacing components of $\mathfrak{x}_n$ by solutions of their variational equations. This replacement is not accomplished for all components of $\mathfrak{x}_n$ simultaneously, but in an alternating procedure: with z_n fixed, x_n and y_n are replaced by x^*_n, y^*_n satisfying $\mathcal{E}(x^*_n;\zeta_n)$, $\mathcal{E}(y^*_n;\zeta_n)$; then, with x^*_n and y^*_n fixed, z_n is replaced by $z^*_n(=\exp\zeta^*_n)$ satisfying $\mathcal{E}(\zeta^*_n;x^*_n,y^*_n)$. On $C^*(O,1)$, $x^*_n=x_n$, $y^*_n=y_n$, $z^*_n=z_n$. We shall show in this section that these "Dirichlet problem" solutions x^*_n, y^*_n, z^*_n exist; are unique; are bounded above (excepting z^*_n) and below, independently of n; are analytic in $C(O,1)$ and analytically extensible outside $\bar{C}(O,1)$; and decrease the B-integral, so that if $\mathfrak{x}^*_n - x^*_n, y^*_n, z^*_n$, $B[\mathfrak{x}^*_n] < B[\mathfrak{x}_n]$, unless $\mathfrak{x}^*_n \equiv \mathfrak{x}_n$. Since values on the boundary are unchanged, the new vectors are equicontinuous there. From the definition of a "minimal sequence in approximation," $B[\mathfrak{x}^*_n] \geqq \beta$, so that $\{\mathfrak{x}^*_n\}$ is also a minimal sequence, whose components satisfy the equations $\mathcal{E}(x^*_n;\zeta_n)$, $\mathcal{E}(y^*_n;\zeta_n)$, $\mathcal{E}(\zeta^*_n;x^*_n,y^*_n)$.

We shall assume, in accordance with § 4, that all our data are analytic, and shall make no attempt to determine minimum restrictions in the following statements. Many are obviously true under weaker hypotheses.

Noting that $\mathcal{E}(x;\zeta)$, $\mathcal{E}(y;\zeta)$ are the same, we may consider just the equation $\mathcal{E}(x;\zeta)$:

$$\Delta x - 2(\zeta_u x_u + \zeta_v x_v) = 0.$$

Our problem is: given that ζ is analytic in $C(O,1+\epsilon)$, where $\epsilon>0$, a function x^* satisfying $\mathcal{E}(x^*;\zeta)$ is to be found, of class C'' in $C(O,1)$, continuous on $\bar{C}(O,1)$, and with pre-assigned analytic boundary values.

Analyticity of ζ in $C(O,1+\epsilon)$ assures existence of the Green's function for $\mathcal{E}(x;\zeta)$ which vanishes on $C^*(O,1)$. Hence existence of a solution x^* follows by the generalized Poisson formula.[11] A non-constant solution of $\mathcal{E}(x;\zeta)$ can have no maximum or minimum at an interior point.[12] Consequently, a solution x^*_n of the boundary value problem for $\mathcal{E}(x;\zeta_n)$ is unique, and possesses upper and lower bounds which depend only on the boundary values—i. e., on Γ alone, not on n. Thus x^*_n, y^*_n are bounded uniformly.

---

[11] D. Hilbert, *Grundzüge einer allgemeinen Theorie der Integralgleichungen*, Leipzig, 1912, p. 243.

[12] Courant-Hilbert II, p. 275.

We may write $\mathcal{E}(\zeta; x, y)$ in the form

$$\Delta\zeta + p(u,v)\exp(-2\zeta) = 0, \tag{13}$$

where $p(u,v) \geqq 0$, being a sum of squares of derivatives of $x(u,v)$, $y(u,v)$, and is analytic in $C(O, 1+\epsilon)$. Existence of a function ζ^* of class C'' satisfying (13) and assuming (continuously) prescribed analytic boundary values may be proved in various ways. For instance, if one writes $\zeta^{(h)}$ for the harmonic function coinciding on $C^*(O,1)$ with ζ^*, the function $\bar{\zeta} = 2(\zeta^{(h)} - \zeta^*)$ must be found so as to satisfy

$$\Delta\bar{\zeta} - w\bar{\zeta} = w(\exp(\bar{\zeta}) - \bar{\zeta}),$$

where the analytic function $w(u,v) \geqq 0$. Bieberbach has proved, in solving $\Delta z = \exp(z)$, the convergence of a sequence of successive approximations for this equation.[13]

The solution is unique; to prove this suppose that ζ_1, ζ_2 are solutions of (13) which coincide on $C^*(O,1)$. By the theorem of the mean of differential calculus,

$$\Delta\zeta_1 - \Delta\zeta_2 = -p\{\exp(-2\zeta_1) - \exp(-2\zeta_2)\} = 2p(\zeta_1 - \zeta_2)\exp(-2\bar{\zeta}),$$

where $\bar{\zeta}$ is analytic and lies between ζ_1 and ζ_2. The difference $z = \zeta_1 - \zeta_2$ satisfies the equation

$$\Delta z - 2p\exp(-2\bar{\zeta})z = 0$$

and vanishes on $C^*(O,1)$; it consequently vanishes identically.[14]

From the totally geodesic minimal surfaces mentioned in § 3, one might guess that ζ^* would have its minimum on $C^*(O,1)$, but not necessarily its maximum. This is so. Writing P for (u,v), Q for (u',v'), $\zeta^{(h)}(P)$ as before for the harmonic function which coincides on $C^*(O,1)$ with $\zeta(P)$, and $G(P,Q)$ for the non-negative harmonic Green's function which vanishes on $C^*(O,1)$, we have[15]

$$\zeta^*(P) = \zeta^{(h)}(P) + \int\int G(P,Q)p(Q)\exp(-2\zeta^*(Q))dQ,$$

showing that $\zeta^*(P) \geqq \zeta^{(h)}(P)$. Hence the minimum of ζ^* is attained on the boundary; the maximum may lie in the interior, as is shown by the example of the geodesic sphere. Thus the functions $\{z^*_n\}$ have a uniform, positive lower bound: $z^*_n > m/2$, which depends only on Γ, and not on n. But we know no uniform upper bound.

---

[13] Courant-Hilbert II, p. 286.

[14] Courant-Hilbert II, p. 275

[15] Courant-Hilbert I, p. 316.

Each of the equations $\mathcal{E}(x^*; \zeta_n)$, $\mathcal{E}(y^*; \zeta_n)$, $\mathcal{E}(\zeta^*; x^*_n, y^*_n)$ is of the type

$$\Delta\omega = f(u, v; \omega, \omega_u, \omega_v), \tag{14}$$

where f is an analytic function of its five arguments. Two theorems of S. Bernstein[16] apply to solutions of such equations. We quote them for our special region $C(O, 1)$.

THEOREM 5.1. *If a solution $\omega(u, v)$ of* (14) *is of class C'' in $C(O, 1)$, it is analytic in $C(O, 1)$.*

THEOREM 5.2. *If a solution $\omega(u, v)$ of* (14) *is of class C'' (and consequently analytic) in $C(O, 1)$, is continuous on $\bar{C}(O, 1)$, and reduces on $C^*(O, 1)$ to an analytic function of the arc, it can be continued analytically to the exterior of $\bar{C}(O, 1)$.*

Thus the functions x^*_n, y^*_n, z^*_n are analytic in $C(O, 1)$; and since their boundary values are analytic, they may be extended analytically throughout a circle $C(O, 1 + \eta_n)$, where $\eta_n > 0$. (In fact, the existence of ζ^*_n depends on these theorems.)

It remains to be shown that replacement of x_n, y_n, z_n by x^*_n, y^*_n, z^*_n does not increase the B-integral. To this end, we consider the integral, taken over $C(O, 1)$,

$$I[\omega] = \int\int [\phi(u, v)(\omega_u^2 + \omega_v^2) + \psi(\omega; u, v)]\, du\, dv,$$

in which ϕ, ω, ψ are analytic functions of their respective arguments, and furthermore $\phi \geqq 0$, $\psi_{\omega\omega} \geqq 0$. The integral $I[\omega]$ contains by specialization of ϕ and ω each of the three problems: to minimize $B[\mathfrak{x}]$ with respect to one component, leaving the other two unchanged. The variational equations $(\mathcal{E})$ are contained in (15), which must be satisfied in $C(O, 1)$ by any ω minimizing $I[\omega]$ among functions analytic on $\bar{C}(O, 1)$:

$$(\phi\omega_u)_u + (\phi\omega_v)_v - \tfrac{1}{2}\psi_\omega = 0. \tag{15}$$

Fixing y and ζ in $B[\mathfrak{x}]$, $\phi = \exp(-2\zeta)$ and $\psi_x = 0$, so that we get $\mathcal{E}(x; \zeta)$; similarly we obtain $\mathcal{E}(y; \zeta)$. Fixing x and y, with $\phi = 1$ and $\psi = p \cdot \exp(-2\zeta)$, we get (13) or $\mathcal{E}(\zeta; x, y)$.

---

[16] Theorem 5, 1 may be found, e. g., in S. Bernstein, "Sur la nature analytique des solutions des équations aux dérivées partielles du second ordre," *Mathematische Annalen*, Bd. 59 (1904), pp. 20-76; p. 32. For Theorem 5. 2, see S. Bernstein, "Sur la généralisation du problème de Dirichlet. Première Partie," *Mathematische Annalen*, Bd. 62 (1906), pp. 253-271; p. 254.

Since ϕ and $\psi_{\omega\omega}$ are non-negative, the uniqueness of the solution of (15) may be proved exactly as for (13).

We can now establish

LEMMA 5.1. *If* $\omega(u,v)$ *satisfies* (15), *and if* $\tilde{\omega}(u,v)$ *is any other function analytic on* $\bar{C}(O,1)$ *which coincides with* $\omega(u,v)$ *on* $C^*(O,1)$, $I[\tilde{\omega}] > I[\omega]$.

Write $\tilde{\omega} = \omega + \eta$, so that $\eta(u,v)$ vanishes on $C^*(O,1)$ and is analytic on $\bar{C}(O,1)$. Then

$$(16)\quad I[\tilde{\omega}] = I[\omega+\eta] = \iint \{\phi(\omega_u^2+\omega_v^2) + \psi(\omega+\eta;u,v)\}\,du\,dv \\ + 2\iint \phi(\omega_u\eta_u + \omega_v\eta_v)\,du\,dv + \iint \phi(\eta_u^2+\eta_v^2)\,du\,dv.$$

By application of Taylor's theorem,

$$(17)\quad \psi(\omega+\eta;u,v) = \psi(\omega;u,v) + \eta\psi_\omega(\omega;u,v) + \eta^2\psi_{\omega\omega}(\omega+\theta\eta;u,v),$$

where $0 \leqq \theta \leqq 1$, and $\theta(u,v)$ is a continuous function of (u,v). By partial integration (Green's lemma[17]), which may be extended to the whole of $\bar{C}(O,1)$,

$$(18)\quad \iint (\phi\omega_u\eta_u + \phi\omega_v\eta_v)\,du\,dv = -\iint \eta\{(\phi\omega_u)_u + (\phi\omega_v)_v\}\,du\,dv$$

since $\eta \equiv 0$ on $C^*(O,1)$. Substituting from (17) and (18) into (16),

$$I[\tilde{\omega}] = \iint \{\phi(\omega_u^2+\omega_v^2) + \psi(\omega;u,v)\}\,du\,dv \\ -2\iint \eta\{(\phi\omega_u)_u + (\phi\omega_v)_v - \tfrac{1}{2}\psi_\omega(\omega;u,v)\}\,du\,dv \\ + \iint \eta^2\psi_{\omega\omega}(\omega+\theta\eta;u,v)\,du\,dv + \iint \phi(\eta_u^2+\eta_v^2)\,du\,dv.$$

Since (15) holds in $C(O,1)$, $I[\tilde{\omega}] = I[\omega] + P[\eta]$, where $P[\eta] > 0$ unless $\eta \equiv 0$; i. e., unless $\tilde{\omega} = \omega$.

The claims made for $\{\mathfrak{x}^*_n\}$ have now been justified. Re-naming our minimal sequence $\{\mathfrak{x}_n\}$, we may evidently repeat the process verbatim, and any number of times.

6. Growth of the Dirichlet integral. We have a minimal sequence which, by Courant's selection theorem, converges uniformly on $C^*(O,1)$. Our present goal is to show that after a few repetitions of the replacement procedure of § 5, we arrive at a minimal sequence which is uniformly bounded (with respect to n), together with its partial derivatives of the first three orders, in any interior circle $C(O, 1-R)$. Ascoli's theorem will then be applied to pick a subsequence converging to a limit vector $\mathfrak{x}$ of class C'' in $C(O, 1)$, and this $\mathfrak{x}$ will be shown to solve our Plateau problem.

---

[17] Hurwitz-Courant, *Funktionentheorie*, Berlin (1929), p. 450.

The keys to determining these uniform bounds are (1) a result which is implicit in Morrey's work on quasi-linear elliptic equations,[18] and (2) an extension of that result, together with methods which are classic in the theory of elliptic equations. Both the theorems mentioned are concerned with the "growth" of the generalized Dirichlet integral

$$J(P, \rho) = \int\int g(u, v)(\xi_u^2 + \xi_v^2)\, du\, dv \tag{19}$$

under the hypothesis that $\xi(u, v)$ satisfies

$$(g\xi_u)_u + (g\xi_v)_v = 0. \tag{20}$$

We include a proof of the first as well as of the second of these two theorems, making the assumptions peculiar to our problem.

THEOREM 6.1. (Morrey). *Suppose that g is analytic on $\bar{C}(O, 1)$, and that $0 < \mathfrak{m} < g < \mathfrak{M}$. Suppose further that $P \epsilon C(O, 1-R)$, where $R < \frac{1}{2}$, and that $0 < r < R$. Then if $\xi(u, v)$ satisfies* (20),

$$J(P, r) \leqq J(O, 1)(r/R)^\alpha, \tag{21}$$

where $\alpha = \alpha(\mathfrak{m}, \mathfrak{M}) > 0$.

By (20), there exists a function $\omega(u, v)$ such that

$$g\xi_u = \omega_v, \quad g\xi_v = -\omega_u. \tag{22}$$

Hence

$$J(P, \rho) = \int\int_{C(P,\rho)} (\omega_v\xi_u - \omega_u\xi_v)\, du\, dv = A,$$

the area of the region Ω in the (ω, ξ) plane into which $C(P, \rho)$ is carried. If L is the perimeter of Ω and $s = \rho\theta$ the arc on $C^*(P, \rho)$,

$$L = \int_0^{2\pi} (\omega_s^2 + \xi_s^2)^{\frac{1}{2}} \rho\, d\theta;$$

using the isoperimetric inequality $A \leqq L^2/4\pi$ and Schwarz's inequality,

$$J(P, \rho) \leqq (\rho/2) \int_0^{2\pi} (\omega_s^2 + \xi_s^2) \rho\, d\theta.$$

---

[18] C. B. Morrey, "On the solutions of quasi-linear elliptic partial differential equations," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 126-166, Theorem 2, p. 138, and Lemma 3, p. 140.

Now

$$\omega_s^2 + \xi_s^2 \leqq (g + 1/g)g(\xi_u^2 + \xi_v^2) \leqq (2/\alpha)g(\xi_u^2 + \xi_v^2),$$

where $\alpha = \min(\mathfrak{m}, 1/\mathfrak{M})$. Hence

$$J(P, \rho) \leqq (\rho/\alpha) \int_0^{2\pi} g(\xi_u^2 + \xi_v^2)\rho d\theta,$$

from which $\alpha d\rho/\rho \leqq dJ/J$ so that, by integration from $\rho = r$ to $\rho = R$, $J(P, r) \leqq J(P, R)(r/R)^\alpha$, from which (21) follows.

The extension of Morrey's theorem which we shall need requires that $g(u, v) = g(P)$ satisfy a "a Hölder condition with constant C and exponent β" in $C(O, 1 - R)$, where now $R < 1/4$: i. e., if P_1, P_2, are any two points in $C(O, 1 - R)$,

(23) $$|g(P_1) - g(P_2)| < C(P_1P_2)^\beta,$$

C and β being positive constants associated with $C(O, 1 - R)$.

THEOREM 6.2. *Suppose that $g(P)$, in addition to possessing the properties assumed in Theorem* 6.1, *satisfies in $C(O, 1 - R)$ a Hölder condition* (23) *with constant C and exponent β. Then, for $\xi(u, v)$ satisfying* (20), *and for $P \epsilon C(O, 1 - 2R)$,*

$$J(P, r) < K(r/R)^2,$$

where K is determined by C, β, $\mathfrak{m}$, $\mathfrak{M}$, and $J(O, 1)$.

Writing $g_0 = g(P)$ and using (22), we see first that $J(P, \rho) = A_0/g_0$, where A_0 is the area of the region Ω_0 in the $(g_0\xi, \omega)$ plane into which $C(P, \rho)$ is mapped. By the isoperimetric and the Schwarz inequalities, as before,

(24) $$J(P, \rho) \leqq (\rho/2g_0) \int_0^{2\pi} (g_0^2\xi_s^2 + \omega_s^2)\rho d\theta.$$

Now

$$\begin{aligned}(1/g_0)(g_0^2\xi_s^2 + \omega_s^2) &\equiv (g_0/g)g(\xi_u^2 + \xi_v^2) \\ &\quad + [(g^2 - g_0^2)/gg_0]g(\xi_u v_s - \xi_v u_s)^2 \\ &< (1 + K_1\rho^\beta)g(\xi_u^2 + \xi_v^2),\end{aligned}$$

where $K_1 = C\mathfrak{m}^{-2}(\mathfrak{m} + 2\mathfrak{M})$. The inequality (24) becomes

$$J(P, \rho) < \tfrac{1}{2}\rho(1 + K_1\rho^\beta)\, dJ/d\rho.$$

Integrating as in Theorem 6.1 from $\rho = r$ to $\rho = R$,

$$J(P, r) < J(P, R)\left(\frac{1 + K_1R^\beta}{1 + K_1r^\beta}\right)(r/R)^2 < K(r/R)^2,$$

with $K = J(O, 1)(1 + K_1)$.

7. Uniform bounds and Hölder conditions for ζ_n. As in § 5, let $\mathfrak{x}_n$: (x_n, y_n, z_n) be the bounded analytic minimal sequence obtained in § 4, and $(x^*_n, y^*_n, \zeta^*_n)$ the solutions of $\mathcal{E}(x^*_n;\zeta_n)$, $\mathcal{E}(y^*_n;\zeta_n)$, $\mathcal{E}(\zeta^*_n;x^*_n,y^*_n)$ having the same boundary values as (x_n, y_n, ζ_n) respectively. We know of x^*_n and y^*_n that they are bounded above and below independently of n, but of ζ^*_n only that they are so bounded below. We shall show, however, that the $\{\zeta^*_n\}$ are uniformly bounded in any interior circle $C(O,1-R)$; once for all, $0<R<1/8$:

1. if $P\in C(O,1-R), \zeta^*_n(P) < H(\Gamma,R)$.

We shall show further that the $\{\zeta^*_n\}$ satisfy a uniform Hölder condition:

2. if $P_1, P_2\in C(O,1-R)$,
$$|\,\zeta^*_n(P_1)-\zeta^*_n(P_2)\,| < L(P_1P_2)^\eta,$$

where $L = L(\Gamma,R)$ and $\eta=\eta(\Gamma,R)>0$.

It is of the greatest importance in the sequel that H, L and η are independent of n.

With an obvious modification of the notation in (13), we may write $\mathcal{E}(\zeta^*_n;x^*_n,y^*_n)$ as $\Delta\zeta^*_n + p^*_n\exp(-2\zeta^*_n) = 0$. We have for ζ^*_n a Green's function representation:

$$(25)\qquad \zeta^*_n(P) = \zeta_n^{(h)}(P) + \iint\limits_{C(O,1)} G(P,Q)p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ.$$

Now

$$(26)\qquad G(P,Q) = -(1/2\pi)\log PQ + \gamma(P,Q),$$

where $\gamma(P,Q)=\gamma(u,v;u',v')$ is harmonic (in u, v for fixed u', v' and conversely) and regular in $C(O,1)$. Since $P\in C(O,1-R)$, the harmonic functions $\zeta_n^{(h)}(P)$ and $\gamma(P,Q)$ are bounded independently of n:

$$(27)\qquad |\,\zeta_n^{(h)}P\,| < H_1(\Gamma,R);$$

$$(28)\qquad |\,\gamma(P,Q)\,| < H_2(R).$$

Since $\{\mathfrak{x}^*_n\}$ is a minimal sequence, we have $B[\mathfrak{x}^*_n] < M(\Gamma)$, so that

$$(29)\qquad \iint\limits_{C(O,1)} p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ < M(\Gamma).$$

Further, for Q outside $C(P,R)$,

$$(30)\qquad -\log PQ < -\log R.$$

Introduce (26) into (25), and break the integral involving $\log PQ$ into two parts: one over $C(P,R)$, the other over the rest of $C(O,1)$. Taking absolute values and using the inequalities (27)-(30), we obtain

(31) $$| \zeta^*_n(P) | < H_1 - (M/2\pi) \log R + MH_2 + (1/2\pi) I_n(R),$$
where

$$I_n(R) = -\iint\limits_{C(P,R)} \log PQ \cdot p^*_n(Q) \exp(-2\zeta^*_n(Q))\, dQ.$$

We shall show that $I_n(R)$ is also bounded, using the theorems of § 6. In terms of polar coördinates $\rho, \theta : \rho = PQ$; $\sin\theta = (v' - v)/PQ$, $\cos\theta = (u' - u)/PQ$, and with $R > \delta > 0$, we may write

$$I_n(R) = \lim_{\delta \to 0} I_n(R, \delta),$$

where

$$I_n(R, \delta) = -\int_0^{2\pi}\int_\delta^R \log \rho \cdot p^*_n \exp(-2\zeta^*_n)\, \rho\, d\rho\, d\theta$$

Integrating by parts,

(32) $$I_n(R, \delta) = I'_n(R, \delta) + I''_n(R, \delta),$$

where

(33) $$I'_n(R, \delta) = -\left[\log \rho \int_0^{2\pi}\int_0^\rho p^*_n \exp(-2\zeta^*_n)\, r\, dr\, d\theta\right]_{\rho=\delta}^{\rho=R}$$

and

(34) $$I''_n(R, \delta) = \int_\delta^R (d\rho/\rho) \int_0^{2\pi}\int_0^\rho p^*_n \exp(-2\zeta^*_n)\, r\, dr\, d\theta.$$

Since by (11)

$$\exp(-2\zeta^*_n) = (z^*_n)^{-2} < (m/2)^{-2},$$

we have, from the definition of p^*_n,

(35) $$\int_0^{2\pi}\int_0^\rho p^*_n \exp(-2\zeta^*_n)\, r\, dr\, d\theta < (m/2)^{-2}\{D[x^*_n \mid C(P, \rho)] + D[y^*_n \mid C(P, \rho)]\},$$

where $D[x^*_n \mid C(P, \rho)]$, for instance, is the ordinary Dirichlet integral of x^*_n taken over $C(P, \rho)$. The analytic z_n which was replaced by z^*_n was less than $(\mathfrak{R} + m)$ by (12). Consequently,

(36) $$D[x^*_n \mid C(P, \rho)] + D[y^*_n \mid C(P, \rho)] < (\mathfrak{R} + m)^2 \int_0^{2\pi}\int_0^\rho p^*_n \exp(-2\zeta_n)\, r\, dr\, d\theta.$$

The integral on the right in (36) is the sum of two integrals of the form

$J(P, \rho)$—cf. (19); x^*_n and y^*_n are solutions of the corresponding variational equations of type (20), with $g = \exp(-2\zeta_n)$, namely of $\mathcal{E}(x^*_n; \zeta_n)$ and $\mathcal{E}(y^*_n; \zeta_n)$, respectively. Since $0 < (\mathfrak{R} + m)^{-2} < \exp(-2\zeta_n) < (m/2)^{-2}$, we have on application of Theorem 6.1 the inequality

$$(37) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta_n) r\,dr\,d\theta < 2M \cdot (\rho/R)^{\alpha},$$

M being as before the uniform bound for $B[\mathfrak{x}_n]$, and α a positive number depending on m and $\mathfrak{R}$—the bounds for Γ—alone. Combining (35), (36) and (37),

$$(38) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta^*_n) r\,dr\,d\theta < H_3(\Gamma, R) \cdot \rho^{\alpha}.$$

Applying (38) to (33) and to (34),

$$I'_n(R, \delta) < -H_3 \cdot (R^{\alpha} \log R + \delta^{\alpha} \log \delta),$$

$$I''_n(R, \delta) < H_3 \int_{\delta}^{R} \rho^{\alpha-1} d\rho = (H_3/\alpha)(R^{\alpha} - \delta^{\alpha}).$$

Returning to (32) and letting δ tend to zero,

$$(39) \qquad I_n(R) = \lim_{\delta \to 0} I_n(R, \delta) < H_3 R^{\alpha}(-\log R + 1/\alpha).$$

With (39) we can complete (31), which we write briefly as

$$(40) \qquad |\, \zeta^*_n(P) \,| < H = H(\Gamma, R),$$

for any point P in $C(O, 1 - R)$.

We now proceed to prove the second assertion of this section: the $\{\zeta^*_n\}$ satisfy a uniform Hölder condition in $C(O, 1 - R)$. Let P_1, P_2 be any two points in $C(O, 1 - R)$. By (25) and (26),

$$(41) \quad \zeta^*_n(P_1) - \zeta^*_n(P_2) = \zeta_n^{(k)}(P_1) - \zeta_n^{(k)}(P_2)$$
$$+ (1/2\pi) \iint_{C(O,1)} \log \{P_2Q/P_1Q\} \cdot p^*_n(Q) \exp(-2\zeta^*_n(Q))\,dQ$$
$$+ \iint_{C(O,1)} \{\gamma(P_1, Q) - \gamma(P_2, Q)\} p^*_n(Q) \exp(-2\zeta^*_n(Q))\,dQ.$$

Since $\zeta_n^{(k)}(u, v)$ and $\gamma(u, v; u', v')$ are harmonic functions of u and v, they satisfy Lipschitz conditions[10] in $C(O, 1 - R)$:

---

[10] O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 227.

(42) $$|\zeta_n^{(k)}(P_1) - \zeta_n^{(k)}(P_2)| < L_1 \cdot (P_1P_2),$$

and, for every Q in $C(O,1)$,

(43) $$|\gamma(P_1,Q) - \gamma(P_2,Q)| < L_2 \cdot (P_1P_2),$$

where $L_1 = L_1(\Gamma, R)$, $L_2 = L_2(R)$. As in (31), we consider absolute values in (41); (42) and (43) yield the inequality

(44) $$|\zeta^*_n(P_1) - \zeta^*_n(P_2)| < (L_1 + ML_2) \cdot (P_1P_2) + (1/2\pi)\left|\iint\limits_{C(O,1)} \log\{P_2Q/P_1Q\} \cdot p^*_n(Q) \exp(-2\zeta^*_n(Q))\,dQ\right|$$

Writing $\psi = \angle P_2P_1Q$ and using the cosine law,

(45) $$\log\{P_2Q/P_1Q\} = \tfrac{1}{2}\log\{1 + (P_1P_2/P_1Q)^2 - 2(P_1P_2/P_1Q)\cos\psi\}.$$

For Q outside $C(P_1, \sqrt{P_1P_2})$, the argument of the logarithm on the right in (45) is greater than $1 + (P_1P_2/2)^2 - P_1P_2$, and at most equal to $1 + P_1P_2 + 2\sqrt{P_1P_2}$. Using the elementary inequality, good for $|x| < a < 1$,

$$|\log(1+x)| < |\log(1-a)|\, a^{-1}\, |x|,$$

we conclude that a constant L_3 exists so that for $P_1Q > \sqrt{P_1P_2}$ with P_1P_2 sufficiently small

(46) $$|\log\{P_2Q/P_1Q\}| < L_3\sqrt{P_1P_2}.$$

Consequently that portion of the integral in (44) which is taken over $C(O,1) - C(P_1, \sqrt{P_1P_2})$ is bounded by $ML_3\sqrt{P_1P_2}$. As to the rest of the integral,

(47) $$\left|\iint\limits_{C(P_1,\sqrt{P_1P_2})} \log\{P_2Q/P_1Q\} p^*_n(Q) \exp(-2\zeta^*_n(Q))\,dQ\right| < \left|\iint\limits_{C(P_1,\sqrt{P_1P_2})} \log P_1Q \cdot p^*_n(Q) \exp(-2\zeta^*_n(Q))\,dQ\right| + \left|\iint\limits_{C(P_2, 2\sqrt{P_1P_2})} \log P_2Q \cdot p^*_n(Q) \exp(-2\zeta^*_n(Q)\,dQ\right|.$$

To bound the two integrals on the right, we note that the proof of (39) establishes the fact that

(48) $$I_n(\rho) < H_3(\Gamma, R)\rho^\alpha(-\log\rho + 1/\alpha)$$

for any $\rho < R$. Since $-\rho^\alpha \log\rho < L_4(\alpha)\rho^{\alpha/2}$, we may write

(49) $$I_n(\rho) < H_3(L_4 + 1/\alpha)\rho^{\alpha/2} = L_5(\Gamma, R)\rho^{\alpha/2}.$$

We now apply (49) to (47), choosing appropriate polar coördinates, and get

$$(50)\qquad \left|\iint_{C(P_1,\sqrt{P_1P_2})} \log\{P_2Q/P_1Q\}\,p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ\right| < L_6(\Gamma,R)\,(P_1P_2)^{\alpha/4}$$

Returning to (44) with inequalities (46) and (50), we obtain

$$(51)\qquad |\,\zeta^*_n(P_1)-\zeta^*_n(P_2)\,| < L(P_1P_2)^{\eta},$$

where $L = L(\Gamma,R)$ and $\eta = \alpha/4 = \eta(\Gamma)$, if P_1, P_2 are any two points of $C(O,1-R)$.

As usual, we re-name our minimal sequence $\{\mathfrak{z}_n\}$.

8. An $\{\mathfrak{z}_n\}$ with uniformly bounded first derivatives. We replace the $\{\mathfrak{z}_n\}$ of § 7 by a new $\{\mathfrak{z}^*_n\}$, as described in § 5. The Green's function representation (25) now holds for the new ζ^*_n. Differentiation of (26) with respect to u leads to the inequality

$$(52)\qquad |\,G_u(P,Q)\,| < (PQ)^{-1} + |\,\gamma_u(P,Q)\,|.$$

Since ζ^*_n, p^*_n and γ are analytic in a circle which contains $\bar{C}(O,1)$, (52) assures the equality

$$(53)\qquad \zeta^*_{nu}(P) = \zeta_{nu}^{(h)}(P) + \iint_{C(O,1)} G_u(P,Q)\,p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ.$$

Suppose that $P \in C(O,1-2R)$. By (42) and (43),

$$(54)\qquad |\,\zeta_{nu}^{(h)}(P)\,| < L_1,$$

$$(55)\qquad |\,\gamma_u(P,Q)\,| < L_2.$$

Taking absolute values in (53) and using (52), (54) and (55):

$$(56)\qquad |\,\zeta^*_{nu}(P)\,| < L_1 + ML_2 + \iint_{C(O,1)} (PQ)^{-1}p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ.$$

Integrating over $C(O,1) - C(P,R)$, we obtain

$$(57)\qquad \iint (PQ)^{-1}p^*_n(Q)\exp(-2\zeta^*_n(Q))\,dQ < M/R.$$

We must now deal with

$$(58)\qquad J_n(R) = \iint_{C(P,R)} (PQ)^{-1}p^*(Q)\exp(-2\zeta^*_n(Q))\,dQ.$$

Integrating by parts in terms of polar coördinates ρ, θ at P, from $\rho = \delta$ to $\rho = R$, we have, as in (32),

$$(59) \qquad J_n(R, \delta) = J'_n(R, \delta) + J''_n(R, \delta),$$

where

$$(60) \qquad J'_n(R, \delta) = [\rho^{-1} \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta^*_n) r dr d\theta]_{\rho=\delta}^{\rho=R},$$

and

$$(61) \qquad J''_n(R, \delta) = \int_\delta^R \rho^{-2} d\rho \cdot \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta^*_n) r dr d\theta.$$

Following the developments of (35)-(37), we must use somewhat different bounds for our present ζ_n, which is that obtained in § 7: we have

$$(62) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta^*_n) r dr d\theta < (m/2)^2 \int_0^{2\pi}\int_0^{\rho} p^*_n r dr d\theta,$$

and by (40)

$$(63) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n r dr d\theta < \exp(2H) \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta_n) r dr d\theta$$

The present ζ_n satisfies the Hölder condition (51) in $C(O, 1-2R)$, so that $\exp(-2\zeta_n)$ also satisfies a uniform Hölder condition. Since equations $\mathcal{E}(x^*_n; \zeta_n)$ and $\mathcal{E}(y^*_n; \zeta_n)$ hold, Theorem 6.2 can be applied:

$$(64) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta_n) r dr d\theta < K(\rho/R)^2,$$

$K(\Gamma, R)$ being independent of n. (62), (63) and (64) yield the inequality

$$(65) \qquad \int_0^{2\pi}\int_0^{\rho} p^*_n \exp(-2\zeta^*_n) r dr d\theta < F_1(\Gamma, R)\rho^2.$$

Returning with (65) to (60) and (61), and letting $\delta \to 0$ in (59), we obtain

$$(66) \qquad J_n(R) < 2RF_1.$$

Returning to (56), we find that

$$(67) \qquad | \zeta^*_{nu}(P) | < F(\Gamma, R),$$

where $F(\Gamma, R) = L_1 + ML_2 + 2RF_1$. The same bound holds for the derivative with respect to v.

We note further that for $P \in C(O, 1 - 2R)$

$$(68) \qquad | \zeta^*_n(P) | < H'(\Gamma, R).$$

The proof of (68) differs from that of (40) in only one particular: it cannot be said of the present ζ_n that, whatever $P \in C(O, 1)$, $\exp(2\zeta_n(P)) < (\mathfrak{R} + m)^2$; this was used in (36). But by (40)—the ζ^*_n there being what are now written ζ_n—$\exp(2\zeta_n(P)) < \exp(2H)$ for all $P \in C(O, 1 - R)$. Accordingly we represent $\zeta^*_n(P)$ in terms of the Green's function $G_R(P, Q)$ which vanishes on $C^*(O, 1 - R)$; thus modified, the reasoning of (25)-(40) leads to (68).

Again we re-name our minimal sequence $\{\mathfrak{r}_n\}$. The next step is to bound the first derivatives of x_n and y_n. We replace the $\mathfrak{r}_n$ just obtained by the vector $\mathfrak{r}^*_n$ derived from $\mathfrak{r}_n$ by § 5. If $G_{2R}(P, Q)$ is the harmonic Green's function which vanishes on $C^*(O, 1 - 2R)$, we have, for $P \in C(O, 1 - 2R)$, the representation

$$(69) \qquad x^*_n(P) = x_n{}^{(h)}(P) - 2 \iint\limits_{C(O,1-2R)} G_{2R}(P, Q) q^*_n(Q)\, dQ;$$

in (69), $x_n{}^{(h)}$ is the harmonic function coinciding with x_n on $C^*(O, 1 - 2R)$, and

$$(70) \qquad q^*_n = \zeta_{nu} x^*_{nu} + \zeta_{nv} x^*_{nv}.$$

In (69), we may differentiate under the integral sign:

$$(71) \qquad x^*_{nu}(P) = x_{nu}{}^{(h)}(P) - 2 \iint\limits_{C(O,1-2R)} [G_{2R}(P, Q)]_u q^*_n(Q)\, dQ.$$

Let us suppose that $P \in C(O, 1 - 3R)$. Then the inequalities

$$(72) \qquad | x_{nu}{}^{(h)}(P) | < L'_1(\Gamma, R),$$

and

$$(73) \qquad | [\gamma_{2R}(P, Q)]_u | < L'_2(\Gamma, R)$$

follow precisely as do (54) and (55). By Schwarz's inequality, we obtain on integrating over $C(O, 1 - 2R)$,

$$(74) \qquad \iint (PQ)^{-1} | q^*_n(Q) |\, dQ$$
$$\leqq [\iint (PQ)^{-1}(\zeta^2_{nu} + \zeta^2_{nv})\, dQ \cdot \iint (PQ)^{-1}(x^*_{nu}{}^2 + x^*_{nv}{}^2)\, dQ]^{\frac{1}{2}}.$$

The bound for derivatives of ζ_n (67) now gives

$$(75) \qquad \iint_{C(O,1-2R)} (PQ)^{-1}(\zeta^2_{nu}+\zeta^2_{nv})\,dQ < 2\pi(1-2R)^2F^2.$$

By (68), $|\zeta_n| < H'$ throughout $C(O,1-2R)$. Thus the second integral on the right in the inequality (74) is bounded by

$$\exp(2H')\iint (PQ)^{-1}(x^*_{nu}{}^2+x^*_{nv}{}^2)\exp(-2\zeta_n)\,dQ.$$

That part which is taken over $C(P,R)$, $P\in C(O,1-3R)$, can be bounded by the procedure used in (58)-(66); while for the rest of the integral, $PQ > R$. Combining these results, we obtain

$$(76) \qquad \iint_{C(O,1-2R)} (PQ)^{-1}(x^*_{nu}{}^2+x^*_{nv}{}^2)\,dQ < \exp(2H')\{2RF'_1+M/R\},$$

F'_1 being a constant analogous to F_1 in (66). Applying the inequalities (72)-(76) to equation (71), we get

$$(77) \qquad |x^*_{nu}(P)| < F'(\Gamma,R)$$

for $P\in C(O,1-3R)$.

It is now readily seen that the new ζ^*_n and its first derivatives are bounded uniformly in $C(O,1-3R)$. We have only to represent $\zeta^*_n(P)$ in terms of the harmonic Green's function $G_{3R}(P,Q)$ which vanishes on $C^*(O,1-3R)$, and to use the boundedness of $p^*_n(P)$ in $C(O,1-3R)$.

Re-naming once more, the new $\{\mathfrak{x}_n\}$ is bounded with its derivatives of first order in the circle $C(O,1-3R)$.

9. Bounds for higher derivatives. Each of the functions $x_n(u,v)$, $y_n(u,v)$, $\zeta_n(u,v)$ satisfies, in $C(O,1)$, an elliptic equation of the form

$$(78) \qquad \Delta\xi = q(u,v),$$

where $q(u,v)$ is a quadratic function of certain first derivatives which are bounded, independently of n, in any circle interior to $C(O,1)$. For a solution ξ of (78) with $|q|$ bounded in, say, $C(O,1-R)$, a theorem of Dini's[20] states that ξ_u, ξ_v satisfy a Hölder condition in $C(O,1-2R)$. In terms of our minimal sequence, the first derivatives of x_n, y_n and ζ_n satisfy a Hölder condition in any interior circle; the constant and the exponent do not depend on n.

---

[20] U. Dini, "Sur la methode des approximations successives pour les équations aux dérivées partielles du deuxieme ordre," *Acta Mathematica*, vol. 25 (1902), pp. 185-230; p. 202.

Since $q(u,v)$ is merely some quadratic combination of partial derivatives which satisfy Hölder conditions, it also satisfies a Hölder condition

Let $P \in C(O, 1-2R)$, and let $G_{2R}(P,Q) = G_{2R}(u,v;u',v')$ be the harmonic Green's function vanishing on $C^*(O, 1-2R)$; if $\xi^{(h)}$ is the harmonic function coinciding with ξ on $C^*(O, 1-2R)$, and if we write

$$G_{2R}(P,Q) = -(1/2\pi)\log PQ + \gamma_{2R}(P,Q),$$

we have

$$(79)\quad \xi_{uu}(P) = \xi_{uu}^{(h)}(P) - (1/2\pi)\frac{\partial^2}{\partial u^2}\int\!\!\int \log PQ\cdot q(Q)\,dQ + \frac{\partial^2}{\partial u^2}\int\!\!\int \gamma_{2R}(P,Q)\,dQ,$$

the integrals being taken over $C(O, 1-2R)$.

The first and third terms on the right in (79) are bounded, by familiar reasoning. If the limit below exists[21]

$$(80)\quad \frac{\partial^2}{\partial u^2}\int\!\!\int_{C(O,1-2R)} \log PQ\cdot q(Q)\,dQ = \lim_{\delta\to 0}\Big[-\int_{C^*(P,\delta)}\frac{\partial}{\partial u}\log PQ\cdot q(Q)\,dv' + \int\!\!\int_{C(O,1-2R)-C(P,\delta)}\frac{\partial^2}{\partial u^2}\log PQ\cdot q(Q)\,dQ\Big]$$

Now

$$(81)\quad \lim_{\delta\to 0}\int_{C^*(P,\delta)}\frac{\partial}{\partial u}\log PQ\cdot q(Q)\,dv' = \pi q(P);$$

while in terms of polar coördinates (r,θ) at P

$$(82)\quad \int\!\!\int\frac{\partial^2}{\partial u^2}\log PQ\cdot q(Q)\,dQ = \int\!\!\int r^{-2}\cos 2\theta[q(Q)-q(P)]\,dQ + \int\!\!\int r^{-2}\cos 2\theta\, q(P)\,dQ.$$

Since for $P, Q \in C(O, 1-2R)$, $|q(Q)-q(P)| < L'r^{\eta}$, we can bound the integrals in (82):

$$(83)\quad (1/2\pi)\int\!\!\int_{C(O,1-2R)}\frac{\partial^2}{\partial u^2}\log PQ\cdot q(Q)\,dQ < L'\eta^{-1}(1-2R)^{\eta} + \max|q|\,(\log 2 - \log R).$$

Applying (81) and (83) to (80), and then going back to (79), we obtain

$$(84)\quad |\xi_{uu}(P)| < L'',$$

---

[21] Courant-Hilbert I, p. 317.

where L'' is, for our minimal sequence, independent of n: the second derivatives of x_n, y_n, ζ_n are uniformly bounded in $C(O, 1-3R)$.

Since $\mathcal{E}(\zeta_n; x_n, y_n)$ holds,

$$\Delta\left(\frac{\partial\zeta_n}{\partial u}\right) = \Phi(\zeta_n; \frac{\partial x_n}{\partial u}, \cdots; \frac{\partial^2 x_n}{\partial u^2}, \cdots), \tag{85}$$

where Φ involves derivatives of at most the second order, and is bounded by (84) in any interior. Dini's theorem can be applied: the second derivatives of ζ_n satisfy a Hölder condition in any interior, with the constant and the exponent independent of n. But we cannot immediately draw the same conclusion for the second derivatives of x_n and y_n: they satisfy equations which involve, not the derivatives of ζ_n, but those of the preceding ζ_n, of whose second derivatives nothing is known. The difficulty is easily overcome: we apply once more the replacement procedure of § 5. All that has so far been obtained carries over to the new $\{\mathfrak{x}_n\}$; we are able in addition to show that all second derivatives satisfy a Hölder condition. By applying the bounding procedure of this section to (85), it is seen that the *third* derivatives are uniformly bounded.

The construction of a minimal sequence with bounded third derivatives practically completes the analytical preparation for convergence. However, in order to assure continuity of the limit vector on the closed circle $\bar{C}(O, 1)$, we shall soon want to know [22] that $D[x_n]$ and $D[y_n]$, the ordinary Dirichlet integrals of x_n and y_n taken over $C(O, 1)$, are uniformly bounded. This will be so if z_n is bounded above, uniformly in n, on $\bar{C}(O, 1)$. It will be recalled that this property of z_n, obtained by Lemma 4. 2, was lost when the replacement procedure of § 5 was used; we shall get it back, while retaining the gains already made.

To bound z_n above, we may make use of either the spherical bounding device of Lemma 4. 2, or the fact that a solution of equation $\mathcal{E}(\lambda^*; \theta)$ attains its maximum and its minimum on the boundary. In either case, we must first transform to spherical coördinates (4), then replace $\lambda_n = \frac{1}{2}\log(x_n^2 + y_n^2 + z_n^2)$ by a function known to be bounded, and finally return to Cartesian coördinates; and we must end up, as we started, with a minimal sequence whose third derivatives are uniformly bounded in any closed interior. The Jacobian of (4),

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = r^2 \sin\theta, \tag{86}$$

is zero only on the z-axis ($\theta = 0$). Inspection of the hyperbolic Dirichlet integral $\mathfrak{B}[\mathfrak{x}]$ and the equations $(\mathcal{E})$ shows that our Plateau problem is unchanged by translation parallel to the plane $z = 0$. We may, hence, suppose Γ, and consequently each of the surfaces defined by $\{\mathfrak{x}_n\}$, to be in the half-plane $x > a > 0$, so that for the corresponding component functions (x_n, y_n, z_n) and (r_n, θ_n, ϕ_n),

---

[22] In order to apply Lebesgue's selection theorem in the concluding section.

$$(87) \qquad \left| \frac{\partial(x_n, y_n, z_n)}{\partial(r_n, \theta_n, \phi_n)} \right| > a^2 > 0$$

everywhere in $\bar{C}(O, 1)$.

Consider a partial derivative of order k of one of the functions $r(u, v)$, $\theta(u, v)$, or $\phi(u, v)$. It is a fraction whose numerator is a polynomial in derivatives of orders $\leqq k$ of $x(u, v)$, $y(u, v)$, $z(u, v)$; $x(r, \theta, \phi)$, $y(r, \theta, \phi)$, $z(r, \theta, \phi)$; and whose denominator is the k-th power of the Jacobian. In an interior circle $C(O, 1 - R)$, the derivatives of x_n, y_n, z_n with respect to u and v are bounded, independently of n. By differentiating (4), it is seen that the derivatives of x, y and z with respect to r, θ and ϕ are bounded if r is bounded: but in $C(O, 1 - R)$, the functions $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ are bounded uniformly, so the same is true of $r_n(u, v)$. Combining these remarks with (87), we can say that the partial derivatives of $r_n(u, v)$, $\theta_n(u, v)$, $\phi_n(u, v)$ are bounded, independently of n, in any interior circle.

We may now replace $\lambda_n = \log r_n$ by λ^*_n, the function which coincides with λ_n on $C^*(O, 1)$ and satisfies $\mathcal{E}(\lambda^*_n; \theta_n)$ in $C(O, 1)$. From the form of $\mathcal{E}(\lambda; \theta)$ it is evident that λ^*_n assumes its extreme values on $C^*(O, 1)$, so that the new r^*_n is uniformly bounded. Further, λ^*_n decreases $B[\mathfrak{x}]$. Finally, it satisfies a linear elliptic equation whose coefficients have bounded second derivatives, so that $\lambda^*_n(u, v)$ has bounded third derivatives: this follows from a theorem of E. Hopf.[23]

The Jacobian of the inverse transformation, being the reciprocal of (86), is now uniformly bounded away from zero on $\bar{C}(O, 1)$. Hence the new $\mathfrak{x}_n(u, v)$:

1. is uniformly bounded on $\bar{C}(O, 1)$;

2. has uniformly bounded partial derivatives of the first three orders on any interior circle.

Finally, we replace x_n, y_n by x^*_n, y^*_n satisfying $\mathcal{E}(x^*_n; \zeta_n)$, $\mathcal{E}(y^*_n; \zeta_n)$, and once more re-name. Properties 1 and 2 above persist.

10. The final minimal sequence, and its convergence. Choose an infinite sequence $\{R_k\}$ of positive radii increasing to unity. In each of the concentric interior circles $C(O, R_k)$ our $\mathfrak{x}_n$ is bounded, together with all its derivatives of the first three orders, by a bound which does not depend on n. Consequently the coördinate functions and their derivatives of the first two orders satisfy the conditions of Ascoli's theorem [24] in each $C(O, R_k)$.

---

[23] E. Hopf, "Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung," *Mathematische Zeitschrift*, Bd. 34 (1932), pp. 194-233; Satz 3, p. 210.

We accordingly pick from $\{\mathfrak{x}_n\}$ a subsequence for which the x_n's converge uniformly in $C(O, R_1)$; from this subsequence we pick one for which the y_n's converge uniformly in $C(O, R_1)$; proceeding in regular succession, we obtain a subsequence $\{\mathfrak{x}_n^{(1)}\}$ of $\{\mathfrak{x}_n\}$ whose coördinate functions converge uniformly, together with their derivatives of first and second orders, in $C(O, R_1)$. From $\{\mathfrak{x}_n^{(1)}\}$ we pick in the same fashion a subsequence $\{\mathfrak{x}_n^{(2)}\}$ which converges uniformly, with its derivatives of first two orders, in $C(O, R_2)$. Repeating for successive R_k, we obtain a sequence of successive subsequences $\{\{\mathfrak{x}_n^{(k)}\}\}$ such that $\{\mathfrak{x}_n^{(k)}\}$ converges uniformly with its first and second order derivatives in $C(O, R_k)$. Hence the diagonal sequence $\{\mathfrak{x}_n^{(n)}\}$—which we designate by $\{\mathfrak{x}_n\}$—is a minimal sequence which converges uniformly, together with its derivatives of first and second orders, in any closed region interior to $C(O, 1)$.

At any point of $C(O, 1)$, the limit vector $\mathfrak{x}(u, v)$ is continuous and possesses continuous derivatives of first and second orders which are limits of the corresponding derivatives of $\mathfrak{x}_n$. By § 4, $\{\mathfrak{x}_n\}$ converges uniformly on $C^*(O, 1)$ to a continuous mapping of $C^*(O, 1)$ onto Γ. But we cannot immediately conclude that $\mathfrak{x}(u, v)$ satisfies the requirements of § 2: $\mathfrak{x}(u, v)$ is not obviously continuous on $\bar{C}(O, 1)$.

We shall base our proof of the continuity of $\mathfrak{x}(u, v)$ in approach to $C^*(O, 1)$ on two lemmas; the first of these enables us to infer that $\mathfrak{x}$ satisfies equations $(\mathcal{E})$ in $C(O, 1)$.

LEMMA 10. 1. *If $\mathfrak{x}$ is the limit vector just obtained, and if $\mathfrak{x}'$*

(i) *is of class C' on $C(O, R)$, $R < 1$;*

(ii) *coincides with $\mathfrak{x}$ on $C^*(O, R)$;*

then

$$B[\mathfrak{x}' \mid C(O, R)] \geqq B[\mathfrak{x} \mid C(O, R)].$$

(In the following proof, all B- and D- integrals over $C(O, R)$ will be indicated as above; if the region is not specified, it is $C(O, 1)$).

Suppose that $\mathfrak{x}'$ is as described, and that a positive α exists such that

$$B[\mathfrak{x}' \mid C(O, R)] < B[\mathfrak{x} \mid C(O, R)] - \alpha. \tag{88}$$

---

[24] Courant-Hilbert I, p. 49.

If (88) is possible, we shall construct a new minimal sequence $\{\mathfrak{x}''_n\}$ such that

$$\overline{\lim}\, B[\mathfrak{x}''_n] < \beta$$

which is absurd. First we define a sequence $\{\mathfrak{x}'_n\}$:

$$(89) \qquad \begin{cases} \mathfrak{x}'_n = \mathfrak{x}_n \text{ outside } C(O,R); \\ \mathfrak{x}'_n = \mathfrak{x}' + \mathfrak{z}_n \text{ inside } C(O,R); \end{cases}$$

where $\mathfrak{z}_n$ is the vector, harmonic in $C(O,R)$, which equals $\mathfrak{x}_n - \mathfrak{x}'_n$ on $C^*(O,R)$. Since $\mathfrak{x}_n$ and its first derivatives converge uniformly on $C^*(O,R)$ to $\mathfrak{x}$ and its first derivatives, $\mathfrak{z}_n$ and its tangential derivative converge uniformly to zero on $C^*(O,R)$. Further,

$$|\mathfrak{z}_n| < \epsilon_n \to 0$$

everywhere in $\bar{C}(O,R)$. Thus

$$B[\mathfrak{x}'_n \mid C(O,R)] = B[\mathfrak{x}' + \mathfrak{z}_n \mid C(O,R)]$$
$$< \iint\limits_{C(O,R)} \{\mathfrak{x}'^2_u + \mathfrak{x}'^2_v + 2(\mathfrak{x}'_u \mathfrak{z}_{nu} + \mathfrak{x}'_v \mathfrak{z}_{nv}) + \mathfrak{z}_{nu}^2 + \mathfrak{z}_{nv}^2\}\,(\mathfrak{z}' - \epsilon_n)^{-2} dud$$

whence, m' being a (positive) lower bound of $z'(u,v)$,

$$(90) \quad B[\mathfrak{x}'_n \mid C(O,R)] < (1+\eta_n)\{B[\mathfrak{x}' \mid C(O,R)]$$
$$+ 2(m')^{-2}(D[\mathfrak{x}' \mid C(O,R)] \cdot D[\mathfrak{z}_n \mid C(O,\mathrm{R})])^{\frac{1}{2}} + (m')^{-2} D[\mathfrak{z}_n \mid C(O,R)]\}$$

by the Schwarz inequality; here $\eta_n \to 0$ with ϵ_n.

The harmonic vector function $\mathfrak{z}_n(u,v)$ has a (vectorial) Fourier expansion of the form

$$\mathfrak{z}_n = \mathfrak{a}_{n,0} + \sum_{k=1}^{\infty} (\mathfrak{a}_{n,k} \cos k\theta + \mathfrak{b}_{n,k} \sin k\theta) r^k,$$

whence

$$(91) \qquad D[\mathfrak{z}_n \mid C(O,R)] = \sum_{k=1}^{\infty} k(\mathfrak{a}^2_{n,k} + \mathfrak{b}^2_{n,k}) R^{2k}.$$

By Parseval's identity,

$$(92) \qquad \int_0^{2\pi} \left(\frac{\partial \mathfrak{z}_n}{R\partial\theta}\right)^2 d\theta = \sum_{k=1}^{\infty} k^2 (\mathfrak{a}^2_{n,k} + \mathfrak{b}^2_{n,k}) R^{2k-2}.$$

Comparison of (91) and (92) yields

$$D[\mathfrak{z}_n \mid C(O,R)] < \int_0^{2\pi} \left(\frac{\partial \mathfrak{z}_n}{R\partial\theta}\right)^2 d\theta;$$

since the tangential derivative of $\mathfrak{z}_n$ on $C^*(O,R)$ converges uniformly to zero,

$$\lim D[\mathfrak{z}_n \mid C(O,R)] = 0.$$

Returning to (90),

$$\overline{\lim}\, B[\mathfrak{x}'_n \mid C(O,R)] \leqq B[\mathfrak{x}' \mid C(O,R)],$$

so that by (88) and (89)

$$\overline{\lim}\, B[\mathfrak{x}'_n] \leqq \beta - \alpha,\ \alpha > 0.$$

This is not quite the contradiction we seek: $\mathfrak{x}'_n$ is not necessarily an admissible vector, since its derivatives may be discontinuous on $C^*(O,R)$. Well known methods permit the replacement of $\mathfrak{x}'_n$ by a "smooth" vector $\mathfrak{x}''_n$: we may, *e. g.*, define

$$\mathfrak{x}''_n(P) = (1/\pi\alpha_n^2) \iint\limits_{C(P,\alpha_n)} \mathfrak{x}'_n(Q)\,dQ$$

where $\{\alpha_n\}$ is a suitably chosen sequence with zero limit; $\{\mathfrak{x}''_n\}$ is a minimal sequence such that $\overline{\lim}\, B[\mathfrak{x}''_n] \leqq \beta - \alpha$, whereas, from the definition of β, $\underline{\lim}\, B[\mathfrak{x}''_n] \geqq \beta$.

The vector $\mathfrak{x}(u,v)$ is of class C'' and minimizes $B[\mathfrak{x} \mid C(O,R)]$. By the classical results of the calculus of variations, $\mathfrak{x}(u,v)$ satisfies equations $(\mathcal{E})$ in $C(O,R)$, and hence in $C(O,1)$.

Since $\mathfrak{x}$ is continuous in $C(O,1)$ and on $C^*(O,1)$, we need only show that it is continuous in radial approach to the boundary, uniformly for all radii. We shall need

LEMMA 10.2. *If* $f(r,\phi)$ *is of class* C' *on* $\bar{C}(O,1)$: $0 \leqq r \leqq 1$, $0 \leqq \phi < 2\pi$; *and if*

$$A_h[f(r,\psi)] = (1/2h) \int_{\psi-h}^{\psi+h} f(r,\phi)\,d\phi,$$

then, for $0 < \delta < \frac{1}{2}$ *and with* $k = (2\log 2)^{\frac{1}{2}}$

$$|\, A_h[f(1,\psi)] - A_h[f(1-\delta,\psi)] \,| < k\,(D[f]/2h)^{\frac{1}{2}}\delta^{\frac{1}{2}}. \tag{93}$$

We first observe that

$$f(1,\psi) - f(1-\delta,\psi) = \int_{1-\delta}^{1} \frac{\partial f}{\partial r}\,dr,$$

whence

$$\int_{\psi-h}^{\psi+h} [f(1,\phi) - f(1-\delta,\phi)]\,d\phi = \int_{\psi-h}^{\psi+h}\int_{1-\delta}^{1} \frac{\partial f}{\partial r}\,dr\,d\phi,$$

or

$$A_h[f(1,\psi)] - A_h[f(1-\delta,\psi)] = (1/2h) \int_{\psi-h}^{\psi+h}\int_{1-\delta}^{1} \frac{\partial f}{\partial r}\, dr d\phi.$$

By Schwarz's inequality,

$$| A_h[f(1,\psi)] - A_h[f(1-\delta,\psi)] | \leqq (1/2h)\{ \int_{\psi-h}^{\psi+h}\int_{1-\delta}^{1} r^{-1} dr d\phi \cdot \int_{\psi-h}^{\psi+h}\int_{1-\delta}^{1} \left(\frac{\partial f}{\partial r}\right)^2 r dr d\phi\}^{\frac{1}{2}}$$

which gives (93).

Now form $A_h[\lambda_n]$, where as usual $\lambda_n = \frac{1}{2}\log(x_n^2 + y_n^2 + z_n^2)$. Since $\lambda_n \rightarrow \lambda$ uniformly in any closed interior and on $C^*(O,1)$, $A_h[\lambda_n] \rightarrow A_h[\lambda]$ uniformly on such sets. Using (93) and the fact that $D[\lambda_n] < B[\mathfrak{x}_n] < M$,

$$| A_h[\lambda_n(1,\psi)] - A_h[\lambda_n(1-\delta,\psi)] | < k\ (M/2h)^{\frac{1}{2}}\delta^{\frac{1}{2}},$$

Keeping h and δ fixed and letting $n \rightarrow \infty$,

$$| A_h[\lambda(1,\psi)] - A_h[\lambda(1-\delta,\psi)] | < k\ (M/2h)^{\frac{1}{2}}\delta^{\frac{1}{2}}.$$

so that $A_h[\lambda]$ is continuous in approach to $C^*(O,1)$, and consequently is continuous on $\bar{C}(O,1)$.

From Lemma 10. 1 we know that $\mathcal{E}(\lambda;\theta)$ holds for $\mathfrak{x}$, throughout $C(O,1)$. Integrating both members of $\mathcal{E}(\lambda;\theta)$ and applying the theorem of the mean of integral calculus, we find that $A_h[\lambda]$ satisfies an equation of the form

$$\Delta z + a(u,v) z_u + b(u,v) z_v = 0;$$

the (continuous) solution $z = A_h[\lambda]$, in any closed subregion of $\bar{C}(O,1)$, assumes its maximum and its minimum values on the boundary (§ 5), so that it is monotone in the sense of Lebesgue. According to a fundamental theorem of Lebesgue,[25] a sequence of continuous monotone functions contains a uniformly convergent subsequence, provided only that the functions converge uniformly on the boundary, and that their Dirichlet integrals be uniformly bounded. Choose a sequence of positive numbers $h_i \rightarrow 0$; then the sequence of averages formed with $\{h_i\}$ converges uniformly on $C^*(O,1)$; each average is, as stated above, monotone and continuous on $\bar{C}(O,1)$; the Dirichlet integrals of the averages are uniformly bounded.[26] By Lebesgue's theorem, there is a subsequence $\{h_i\}$ such that $A_{k_i}[\lambda] \rightarrow \lambda$ uniformly on $\bar{C}(O,1)$. It follows that $\lambda(u,v)$ is continuous on $\bar{C}(O,1)$.

---

[25] H. Lebesgue, "Sur le principe de Dirichlet," *Rendiconti del Circolo Matematico di Palermo,* vol. 24 (1907), pp. 371-402; p. 386.

[26] C. B. Morrey, "A class of representations of manifolds, Part I," *American Journal of Mathematics,* vol. 55 (1933), pp. 683-707; p. 687.

By § 9, the sequences $\{x_n\}$ and $\{y_n\}$ also satisfy the conditions of Lebesgue's theorem. We conclude that their limits $x(u,v)$ and $y(u,v)$ are continuous on $\bar{C}(O,1)$; so also is $z(u,v)$, from the continuity of $\lambda(u,v)$.

We conclude that $B[\mathfrak{x}] = \beta$. For, by the preceding discussion, $\mathfrak{x}(u,v)$ is one of the vectors in competition in the minimum problem: hence $B[\mathfrak{x}] \geqq \beta$. On the other hand (lower semi-continuity of $B[\mathfrak{x}]$),[27] $B[\mathfrak{x}] \leqq \underline{\lim}\, B[\mathfrak{x}_n] = \beta$. Hence $B[\mathfrak{x}] = \beta$. Since $\mathfrak{x}(u,v)$ is an admissible vector spanning Γ, and since for such an $\mathfrak{x}(u,v)$ $B[\mathfrak{x}] \geqq b$, and since further $b \leqq \beta$, $B[\mathfrak{x}] = \beta = b$; so that the vector $\mathfrak{x}(u,v)$ actually solves the minimum problem of § 3; and, consequently, the Plateau problem in hyperbolic space.

We may finally show that the limit mapping of $C^*(O,1)$ onto Γ is not only continuous and monotone, but also $(1-1)$; the proof which follows is exactly like Courant's for the Euclidean problem.[28]

Suppose that over a whole arc c on $C^*(O,1)$, $\mathfrak{x}$ has the constant value $\mathfrak{x}_0$. We map $\bar{C}(O,1)$ conformally onto one of its semicircles, C', so that c goes into the bounding diameter of C'. As was remarked in § 3, the B-integral of the new vector $\mathfrak{x}'$, taken over the semi-circle C', is the same as that of $\mathfrak{x}$, taken over $C(O,1)$. If in the complementary semicircle C'' we define $\mathfrak{x}' = \mathfrak{x}_0$,

$$B[\mathfrak{x}'] = B[\mathfrak{x}] = b.$$

Now $\mathfrak{x}'$ certainly is not a solution of $(\mathcal{E})$ corresponding to its boundary values on $C^*(O,1)$. But, as a special case of our existence proof, we are able to solve the system $(\mathcal{E})$ with the new boundary values, and the solution must have a B-integral smaller than that of $\mathfrak{x}'$; which contradicts the definition of b.

THE UNIVERSITY OF CALIFORNIA, BERKELEY,
IOWA STATE COLLEGE, AMES.

---

[27] R. Courant, *l. c.*[7], p. 692

[28] R. Courant, *l. c.*[7], p. 696.

SPHERICAL THEOREMS OF HELLY TYPE AND CONGRUENCE INDICES OF SPHERICAL CAPS.*[1]

By CHAS. V. ROBINSON.

A metric space M has *congruence order* k [5], with respect to a class $\{S\}$ of semimetric spaces if congruent imbeddability of any member S_0 of $\{S\}$ into M is implied by congruent imbeddability of each k points of S_0 into M. The type of congruence order problem dealt with in this paper is: given the space M and the class $\{S\}$ find the smallest integer k for which M has congruence order k with respect to $\{S\}$. Of particular interest is the case when $\{S\}$ is the class of all subsets of a space containing M congruently. For example, one may seek the congruence order of the ρ-sphere[2] of E_n with respect to subsets of E_n. In spaces which, like the E_n, possess, for each given pair of isometric subsets, a self-isometry carrying one into the other, the idea "B is congruently imbeddable in A" is equivalent to the simpler idea "A will cover B." Hence, the problem of the ρ-sphere can be stated: what is the smallest integer k such that if each k points of a subset S_0 of E_n can be covered by a ρ-sphere, it follows that S_0 can be so covered? Now, since in any semimetric space a subset S_0 is contained in a ρ-sphere if and only if the ρ-spheres with centers in S_0 intersect (any point of the intersection is the center of a ρ-sphere containing S_0), the problem may be stated in the alternate intersectional form: find the smallest integer k such that intersection of each k members of any family of ρ-spheres of E_n implies intersection of all the spheres of the family. The answer to this problem is given directly by an application of the following elegant theorem due to E. Helly [1]: *If each $n+1$ members of a family of bounded, closed, convex subsets of E_n intersect, there is a point common to all the sets.* Applying the theorem to the family of ρ-spheres with center in S_0, we see that if each $n+1$ spheres intersect, then all intersect and hence S_0 can be covered by a ρ-sphere. We have proved as corollary to the Helly theorem that *the ρ-sphere of E_n has congruence order $n+1$ with respect to subsets of E_n.*

The ρ-sphere thus provides an interesting link between the concept of

---

* Received March 3, 1941; Revised August 4, 1941.

[1] Presented to the American Mathematical Society, December, 1940. This paper is a part of a dissertation for the Ph. D. degree, University of Missouri, 1940, and was written under the direction of L. M. Blumenthal.

[2] The solid sphere of E_n with radius ρ. In any semimetric space the ρ-sphere with center p is the set of points x of the space such that $px \leqq \rho$.

congruence order and the Helly theorem in E_n. The problems of the congruence order of the hemisphere and of small spherical caps (ρ-spheres of the spherical surface) led L. M. Blumenthal to pose the question: for what integer k does the intersection of each k members of a family of convex subsets of the n-dimensional sphere imply the existence of a common point? In section **1** of the present paper three answers to this question are given corresponding to restrictions on: (1) a single member of the family, (2) the size of the family, and (3) the diameters of all the sets of the family. These spherical analogues of the Helly theorem have as corollaries the congruence orders of small spherical caps. A necessary lemma to the proof of Theorem 2, which restricts the size of the family, is the proof of the theorem for a family of hemispheres, first found by E. Steinitz [2]. The author gives a geometric proof of this result obtained independently of the Steinitz paper, in which the problem is treated algebraically.

In section **2** a study is made of the congruence orders of small caps with respect to semimetric spaces and one cap is found to display a new congruence order property.

Section 1. Spherical Theorems of Helly Type.

In the following the n-dimensional spherical surface, $|x| = r$ of E_{n+1} will be denoted by $S_{n,r}$.[3] If a half-E_{n+1}, $P(x) \equiv \sum_{i=1}^{n+1} a_i x_i - c \geqq 0$, intersects the $S_{n,r}$, the product is called a *spherical cap* and $P(x) = 0$ the *base plane* of the cap. If $P(o) < 0$ the cap is *small*, and if $P(o) = 0$ it is a hemisphere of the $S_{n,r}$. A set of hemispheres of $S_{n,r}$ is called dependent or independent according as the rank of the matrix of base plane coefficients is less than or equal to the number of hemispheres. Clearly if any lower dimensional plane through the origin lies in the half-space of a hemisphere of the $S_{n,r}$, it must lie in the base plane. Hence if the rank of the base plane matrix of a set of hemispheres of $S_{n,r}$ is k, the set of half-spaces contains a plane of dimension $n + 1 - k$ and no higher, and the hemispheres hence contain an $S_{n-k,r}$ but not an $S_{n+1-k,r}$. Any product of hemispheres of $S_{n,r}$ will be termed a *convex*[4] subset of the $S_{n,r}$ *including* the diametral point-pair $S_{0,r}$.

---

[3] The terms "sphere" and "$S_{2,r}$" are interchangeable.

[4] A more usual definition of convex subset of the sphere is: A closed set A is convex if when $p, q \in A$ and $pq < \pi r$, the minor arc pq is contained in A, and if $pq = \pi r$ *one* semi-great-circle pq is contained in A. All sets convex by this definition are convex by the first definition except the whole sphere, while all sets convex by the first are convex in this sense except the diametral point pair $S_{0,r}$. The theorems which follow can be stated with only slight modification using this latter definition.

LEMMA 1a. *A convex subset of the $S_{n,r}$ which contains no diametral point-pair is contained in a small cap.*

Proof. Since the product of the family of all hemispheres containing the given convex set contains no diametral point-pair, there are $n+1$ independent hemispheres of this family defined by half-spaces $D_i: P_i(x) \geqq 0$ whose base planes have $x = 0$ as their sole common point. The function $P(x) = \sum_{i=1}^{n+1} P_i(x)$ is thus everywhere positive over the product T of these hemispheres and hence takes on a positive minimum m in T. The function $Q(x) = P(x) - m$ is non-negative in T while $Q(o) = -m$ is negative. Hence $Q(x) \geqq 0$ is a half-space defining a small cap containing T and thus the given set.

LEMMA 1b.[5] *If each $n+1$ members of a family of hemispheres of the $S_{n,r}$ have a point in common with the product of $n+1$ fixed independent hemispheres of the $S_{n,r}$, then all the hemispheres have a common point.*

Proof. Let D be the product of the half-spaces of the fixed independent hemispheres, let $T = D \cdot S$, let π be the base plane and $Q(x) \geq 0$ the half-space of a small cap containing T (Lemma 1a), and let H_i be the hemispheres of the family with half-spaces D_i. Consider now the family of subsets $C_i = \pi \cdot D \cdot D_i$ of π. These are closed, convex and bounded ($|x| = r$, $x \in C_i$), and any $n+1$ of them $C_{i_1}, C_{i_2}, \cdots, C_{i_{n+1}}$ intersect. For there is, by hypothesis, a point p of $H_{i_1} \cdot H_{i_2} \cdots H_{i_{n+1}} \cdot T = D_{i_1} \cdot D_{i_2} \cdots D_{i_{n+1}} \cdot D \cdot S_{n,r}$, and since the origin o and the set T are on opposite sides of π, the ray op cuts π in a point p' between o and p. Since the sets D_i and D are convex, p' belongs to each of them and hence to $C_{i_1}, C_{i_2}, \cdots$, and $C_{i_{n+1}}$. We have shown that each $n+1$ of the sets C_i, intersect. By the Helly theorem it follows that there is a point q' of π common to them all. The point q of the $S_{n,r}$ on the ray oq' is common to all the hemispheres H_i.

LEMMA 1c. *If each $n+1+k$ members of a family of hemispheres of $S_{n,r}$ have a point in common with the intersection of a fixed set of $n+1-k$ independent hemispheres ($k = 0, 1, \cdots, n+1$), then all the hemispheres have a point in common.*

Proof. Let $H_1, H_2, \cdots, H_{n+1-k}$ be the fixed independent set of hemispheres. The previous lemma proves the case $k = 0$, the anchor for the following induction. We suppose the lemma true for $k-1$, $k-2, \cdots, 0$,

---

[5] As well as simplifying the former proofs, this lemma, suggested by A. N. Milgram, brings out the algebraic nature of the spherical Helly theorems.

$(k > 0)$ and prove it true for k. Enlarge the set $H_1, H_2, \cdots, H_{n+1-k}$ to make as large an independent set of hemispheres as possible, say $H_1, H_2, \cdots, H_{n+1-r}$, $r \leq k$. If $r = k$ then every hemisphere is dependent on the given fixed set, and their intersection (an $S_{k-1,r}$) is common to all the hemispheres. If, however, $r < k$, any $n + 1 + r$ other hemispheres have a point in common with $H_1 \cdot H_2 \cdots H_{n+1-r}$. For, by hypothesis, any $2n + 2$ hemispheres containing the fixed set intersect. By our inductive assumption it follows that all the hemispheres intersect.

The above lemma is interesting when translated into a theorem on inequalities.

THEOREM 1. *If each $n + k + 2$ members of a family of convex subsets of the $S_{n,r}$ intersect, and if one member contains no $S_{k,r}$, $0 \leq k \leq n$, then there is a point common to all the sets of the family.*

Proof. Let $\{A\}$ be the given family of sets and $\{H\}$ the family of all hemispheres containing sets of $\{A\}$. If A_0 is the set containing no $S_{k,r}$, then there is a set of $n + 1 - k$ independent hemispheres $H_1, H_2, \cdots, H_{n+1-k}$ whose products contains A_0. For if each $n + 1 - k$ hemispheres containing A_0 are dependent, then their base planes contain a common hyperplane of dimension at least $k + 1$ and all the hemispheres containing A_0 have an $S_{k,r}$ in common. Consider a set of $2n + 2$ hemispheres $H_1, H_2, \cdots, H_{2n+2}$ where those beyond $n + 1 - k$ are arbitrary hemispheres of $\{H\}$ which contain respectively sets $A_{n+2-k}, \cdots, A_{2n+2}$ of $\{A\}$. By hypothesis $A_0 \cdot A_{n+2-k} \cdots A_{2n+2} \neq 0$ hence $\prod_{i=1}^{2n+2} H_i \neq 0$. The family $\{H\}$ thus satisfies the conditions of the lemma and possesses a common point p which is a common point of $\{A\}$.

LEMMA 2a. *If q is interior to a convex polyhedron P of the E_n with vertices $p_1, p_2, \cdots, p_m$ $(m \geqq 2n - 1)$ then either (1) q is interior to an n-dimensional sub-polyhedron of P with $2n - 1$ or fewer vertices, or (2) P has exactly $2n$ vertices collinear in pairs with q.*

Proof. The lemma being clearly true for $n = 1$, we make the inductive hypothesis that the lemma holds in the E_s for $s < n$. For $m = 2n - 1$ the lemma is trivial, so we suppose $m \geqq 2n$. Let p'_i be the point in which the ray $p_i q$ meets the boundary of P. Two cases arise:

Case 1. One of the points p'_i is interior to an s-dimensional face P_s of P, $n > s \geqq 1$. By our inductive hypothesis, p'_i is interior to an s-dimensional sub-polyhedron P'_s of P_s which has $2s$ or fewer vertices. The point q is then interior to the $(s+1)$-dimensional polyhedron $C(p_i, P'_s)$[6] of $2s+1 = 2(s+1)-1$ or fewer vertices. If $s+1=n$, the lemma is proved. If, on the contrary, $s+1<n$, then $C(p_i, P'_s)$ cannot be a subset of the boundary $B(P)$ since $q \epsilon C(p_i, P'_s)$, $q \not\epsilon B(P)$. Two $(s+1)$-dimensional faces of P meet in P_s. Taking a vertex not in P_s from each of these faces, say p_j, p_k, q is interior to the $(s+2)$-dimensional polyhedron $C(p_i, p_j, p_k, P'_s)$ of $2s+3 = 2(s+2)-1$ or fewer vertices. If $s+2=n$, the lemma is proved. If $s+2<n$, the process is repeated. It is clear that $n-s$ steps suffice, and that the last step gives an n-dimensional sub-polyhedron of P with $2n-1$ or fewer vertices to which q is interior. Thus for Case 1 the alternative (1) holds.

Case 2. All the points p'_i are vertices of P. Then m is clearly even, and if $m=2n$ we have the alternative (2) of the lemma. If $m \geq 2n+2$, we can choose $2n$ vertices, $p_1, p'_1, p_2, p'_2 \cdots, p_n, p'_n$ collinear in pairs with q, not lying in an $(n-1)$-dimensional hyperplane. A $2n+1$-st vertex p_{n+1} cannot lie in all the n hyperplanes determined by each $n-1$ of the rays qp_i, $i=1, 2, \cdots, n$, since only the point q is common to all these planes, so that p_{n+1} does not lie in one of the hyperplanes, say the plane of $q, p_1, p_2, \cdots, p_{n-1}$. Choose a point $p_{n+2} \neq p'_{n+1}$ on the opposite side of this hyperplane from p_{n+1} (one of the points p_n, p'_n can be used for p_{n+2}). Then q is interior to the n-dimensional polyhedron $C(p_1, p'_1, \cdots, p_{n-1}, p'_{n-1}, p_{n+1}, p_{n+2})$, and the ray $p_{n+1}q$ cannot meet the boundary in a vertex, so that for $m \geq 2n+2$ this case reduces to Case 1.

LEMMA 2b. *A family of more than $2n+2$ hemispheres of $S_{n,r}$ each $2n+1$ of which intersect, have a common point.*

Proof. For $k=n+1$ Lemma 1c states that if each $2n+2$ members of a family of hemispheres intersect then all intersect, so that it suffices here to show that each $2n+2$ intersect. Given any set of $2n+2$ hemispheres $H_1, H_2, \cdots, H_{2n+2}$ another one H_{2n+3} of the family can be added. If the set of spherical centers $p_1, p_2, \cdots, p_{2n+3}$ of these hemispheres lie on a hemisphere with center p, then p is common to the hemispheres H_i, for $pp_i \leq \pi r/2$. Consider the convex polygon P of E_{n+1} with vertices $p_1, p_2, \cdots, p_{2n+3}$. If the p_i's do not lie on a hemisphere then o is interior to P and Lemma 2a applies. Alternative (2) cannot be satisfied, hence (1) holds, i. e. o is interior to a sub-polygon of P with $2n+1$ vertices. This means however that the $2n+1$ hemispheres with these vertices as centers do not intersect, a contradiction to the hypothesis that each $2n+1$ do. Hence each $2n+2$ hemispheres intersect and by Lemma 1c all have a common point.

---

[6] The convex cover of the set $p_i + P'_s$.

THEOREM 2. *A family of more than $2n+2$ convex subsets of $S_{n,r}$ each $2n+1$ of which intersect, has a common point.*

Proof. If one member is not a hemisphere, it contains no $S_{n-1,r}$ and Theorem 1 is satisfied with $k=n-1$. If all the sets are hemispheres the above lemma applies.

Remark. The only type of family of convex subsets of $S_{n,r}$ such that each $2n+1$ intersect while there is no point common to all is a set of $n+1$ independent pairs of opposite hemispheres.

Proof. If one member were not a hemisphere, then by Theorem 1 the sets would have a common point. Hence all the sets are hemispheres and there must not be more than $2n+2$ by Lemma 2b. Since the hemispheres do not intersect, the convex cover of their centers has o as interior point. Since each $2n+1$ hemispheres intersect, o cannot be interior to the convex cover of $2n+1$ centers. Hence the alternative (2) of Lemma 2a holds, which shows that the hemispheres are opposite in pairs.

We state without proof the following separation lemma which is needed in the proof of Theorem 3.

LEMMA 3a. *If A and B are two mutually exclusive convex subsets of the $S_{n,r}$ neither of which contains an $S_{o,r}$ then there is an $S_{n-1,r}$ separating A and B.*

THEOREM 3. *If each $n+1$ members of a family $\{A\}$ of convex subsets of $S_{n,r}$ of arc diameters*[7] *less than $2\pi r/3$ intersect, then all have a common point.*

Proof. The proof is an induction which we anchor with the case $n=1$. Let A_0 be a particular arc of $\{A\}$, c its middle point, c^* the point of the circle opposite c, and T the tangent to the circle at c. Then c^* is not contained in any arc of the family $\{A\}$ and the projection of the family on T from c^* yields a family $\{A'\}$ of segments of T. Each two of these segments have a point in common, and hence, by the Helly theorem, there is a point m' common to all the segments of $\{A'\}$. The corresponding point m of the circle is common to all the arcs of $\{A\}$.

We now suppose the theorem true for the $S_{n-1,r}$ and prove it true for the $S_{n,r}$. If we show that each $n+2$ sets of the family have a common point, the validity of the theorem follows from Theorem 1. Let then $A_0, A_1, \cdots, A_{n+1}$

---

[7] The arc diameter of a set is the maximum arc distance between two points of the set.

be members of $\{A\}$, and suppose that $A_0 \cdot D = 0$, where $D = \prod_{i=1}^{n+1} A_i$. Then there is, by the lemma, an $S_{n-1,r}$, K, separating A_0 and D $(K \cdot A_0 = K \cdot D = 0)$. Each of the $n+1$ sets $D_1 = A_2 \cdot A_3 \cdots A_{n+1}$, $D_2 = A_1 \cdot A_3 \cdots A_{n+1}, \cdots$, $D_{n+1} = A_1 \cdot A_2 \cdots A_n$ contains D and, by hypothesis, a point of A_0, and hence must cross K and have a non-null product $D'_i = K \cdot D_i$ with it. These sets D'_i are the products of n-tuples of the $n+1$ convex subsets $A'_i = K \cdot A_i$, of $S_{n-1,r}$, $i = 1, 2, \cdots, n+1$, of arc diameters less than $2\pi r/3$. By inductive hypothesis there is a point m common to $A'_1, A'_2, \cdots, A'_{n+1}$. But then m belongs to $K \cdot D$, which is consequently not null, a contradiction. Thus each $n+2$ sets of $\{A\}$ have a common point, and the theorem is proved.

Among Theorems 1, 2, 3, Theorem 3 is the truest correspondent to the Helly theorem for the E_n in the sense that its proof follows closely that of the latter, and it shows that locally the E_n and the $S_{n,r}$ have the same Helly property (intersection of each $n+1$ sets of a family of convex sets implies the intersection of all). The diameter bound $2\pi r/3$ of Theorem 3 is not the best bound except for the circle $S_{1,r}$. Theorem 4, as a refinement of Theorem 3 for for the case $n = 2$, finds the largest diameter bound for the sphere.[3]

Preliminary to proving Theorem 4 we state without proof an elementary property of the equilateral spherical triangle and prove two lemmas.

Remark. If a spherical triangle T is of arc diameter less than d, it has a smaller area than the equilateral spherical triangle of arc diameter d.

LEMMA 4a. *If C_1, C_2, C_3, C_4 are convex subsets of the sphere and p_1, p_2, p_3, p_4 are points of a small cap with $p_1 \in C_2 \cdot C_3 \cdot C_4$, $p_2 \in C_1 \cdot C_3 \cdot C_4$, $p_3 \in C_1 \cdot C_2 \cdot C_4$, and $p_4 \in C_1 \cdot C_2 \cdot C_3$, then there is a point common to the sets C_1, C_2, C_3, C_4.*

Proof. If we let π be the base plane of the small cap $\mathcal{K}$ which contains the points p_i, then the projections of the convex sets $C_i \cdot \mathcal{K}$ on π from o are convex domains of π having each three a projection of one p_i in common. By the Helly theorem, there is a point m' common to them all, and the corresponding point m of the sphere is common to the sets C_1, C_2, C_3, C_4.

LEMMA 4b. *If each three points of the subset p_1, p_2, p_3, p_4 of the sphere are on a small cap and the sum of the areas of the four triangles with vertices in the set is less than $4\pi r^2$, then the four points are on a small cap.*

Proof. Let T be the tetrahedron (degenerate or not) of E_3 with vertices p_i. Now o is exterior to T, for if it were interior the sum of the areas of the four triangles would be $4\pi r^2$. Neither can o be in a face of T, since the vertices of any face are in a small cap. Hence, the projection of T on the sphere lies in a small cap.

THEOREM 4. *If each member of a family of convex subsets of the sphere is of diameter less than* $r \cdot \cos^{-1}(-3^{-\frac{1}{2}})$ *and if each three sets have a point in common, then there is a point common to all the sets.*

Proof. Let $\{A\}$ be the given family of sets. If we can show that each four sets of the family have a common point, the validity of the theorem will follow from Theorem 1 with $n = 2$, $k = 0$. Let A_1, A_2, A_3, A_4 be four sets of $\{A\}$ with $p_1 \epsilon A_2 \cdot A_3 \cdot A_4$, $p_2 \epsilon A_1 \cdot A_3 \cdot A_4$, $p_3 \epsilon A_1 \cdot A_2 \cdot A_4$, $p_4 \epsilon A_1 \cdot A_2 \cdot A_3$. Each of the four triangles with vertices in $\{p_i\}$ is contained in one of the sets A_i so that the diameter of the triangle is less than $r \cdot \cos^{-1}(-3^{-\frac{1}{2}})$ the diameter of the equilateral triangle of area πr^2.[8] Hence, the areas of each of these triangles is, by the second remark, less than πr^2, the points p_i are on a small cap by the Lemma 4b, and the four sets A_i have a point in common by Lemma 4a.

COROLLARY. *If no member of a family of convex subsets of the sphere contains a triangle of area* πr^2 *or a diametral point-pair, and if each three members have a common point, there is a point common to all the sets.*

The bound $r \cdot \cos^{-1}(-3^{-\frac{1}{2}})$ of Theorem 4 cannot be increased. For if p_1, p_2, p_3, p_4 be points of the sphere and $p_i p_j = r \cdot \cos^{-1}(-1/3)$, $i \neq j$, then each of the triangles determined is of area πr^2 and diameter $r \cdot \cos^{-1}(-3^{-\frac{1}{2}})$ and the triangles intersect triplewise but have no common point.

It would be desirable to extend Theorem 4 to the $S_{n,r}$, i. e. to enlarge the bound $2\pi r/3$ of Theorem 3 for $n > 2$. The new bound would probably be the arc diameter of a spherical face of an equilateral $n+2$-gon inscribed in the $S_{n,r}$ as it is for $n = 1$ and 2.

Section 2. Congruence Indices for Spherical Caps.

Just as the covering properties of the disc follow from the Helly theorem in the plane, so the covering behavior of spherical caps stems directly from the theorems on the sphere proved in the preceding section.

THEOREM 5. *If each three points of a given subset* S *of the sphere can be covered by a cap of radius* ρ, $\rho < r \cdot \cos^{-1}(1/3)$, *then* S *can be so covered.*

Proof. The triangle of maximum area in a cap is the inscribed equilateral triangle (Remark, p. 266), and the circumradius of the equilateral triangle of area πr^2 is $r \cdot \cos^{-1}(1/3)$, so that a cap of smaller radius contains no triangle of area πr^2. Since each three points of S are by hypothesis on a cap of radius ρ, each three caps of the set of caps of radius ρ and center in S intersect. By the corollary to Theorem 4 there is then a point common to them all and hence S can be covered by a cap of radius ρ.

---

[8] The arc diameter of a spherical triangle may not be equal to a side, e. g. here it is equal to the altitude.

THEOREM 6. *If each four points of a given subset S of the sphere can be covered by a cap of radius ρ, $\rho < \pi r/2$, then S can be so covered.*

The proof follows from Theorem 1 ($n = 2, k = 1$) by an argument similar to the preceding one. Using a similar argument with Lemma 2b we obtain:

THEOREM 7. *If each five points of a subset S of more than six points of the sphere can be covered by a hemisphere, then S can be so covered.*

From Theorem 1 with $n = k = 2$ we obtain:

THEOREM 8. *If each six points of a given set S of the sphere can be covered by a hemisphere, then S can be covered by a hemisphere.*

More general than the covering problems discussed above is the problem of determining the congruence orders of spherical caps with respect to the class of semimetric spaces.

The following concept recently introduced by L. M. Blumenthal [4] will materially facilitate the handling of congruent imbedding properties. A semimetric space S has *congruence indices* $[n, k]$ with respect to a class C of semimetric spaces if the congruent imbedding of any space N of C which contains more than $n + k$ distinct points is implied by the imbedding of each n points (not necessarily distinct) of N in S. The symbols $[n, 0]$ and $[n, 1]$ of S with respect to the class Σ of all semimetric spaces are equivalent respectively to the properties of *congruence order n* and *quasi congruence order n* introduced by K. Menger [3]. It can be seen that if S has the symbol $[n, k]$ it has also any symbol $[n', k']$ for which $n' \geq n$, and $n' + k' \geq n + k$. In the set of symbols which a space has, we introduce a preferential ordering. The symbol $[n, k]$ is preferred to (precedes) $[n', k']$ if $n < n'$, or, in case $n = n'$, if $k < k'$. The first element of the set of preferentially ordered symbols is called the *best* symbol of S with respect to the class C. A *complete set* of symbols for a space is a set which implies all the other symbols which the space has. In terms of congruence indices, Theorems 5, 6, 7 and 8 state that (with respect to the class of subsets of the sphere) a cap of arc radius ρ has congruence symbol $[3, 0]$ for $\rho < r \cdot \cos^{-1}(1/3)$, $[4, 0]$ for $\rho < \pi r/2$, and the symbols $[5, 1]$ and $[6, 0]$

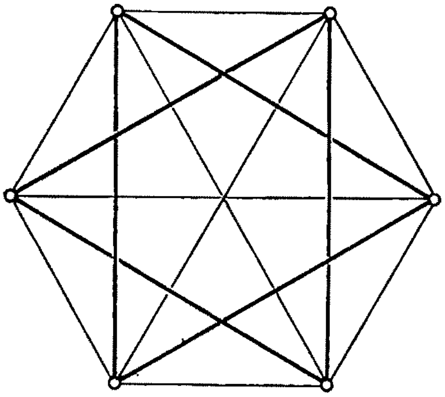
for $\rho = \pi r/2$. It is well known that the sphere has the symbol $[5, 0]$ with respect to Σ.

A *pseudo-spherical set* is one not congruently imbeddable in the sphere while each four of its points are. L. M. Blumenthal has proved the following interesting theorem about such sets ([5], p. 80).

THEOREM. *If P is a pseudo-spherical set containing more than five points, no two of which are diametral, then $pq = r \cdot \cos^{-1}(\pm 1/3)$ for each pair of distinct points p, q of P. At least one distance is $r \cdot \cos^{-1}(-1/3)$.*

An immediate consequence of this theorem is:

THEOREM 9. *The sphere has the congruence symbol $[4, 1]$ with respect to the class of semimetric spaces of diameter less than $r \cdot \cos^{-1}(-1/3)$.*



Proof. If the contrary be supposed, then there exists a semimetric space P of diameter less than $r \cdot \cos^{-1}(-1/3)$ containing more than five points, such that P is not imbeddable in the sphere while each four of its points are. But P is then a pseudo-spherical set satisfying the conditions of Blumenthal's theorem and having no distance equal to $r \cdot \cos^{-1}(-1/3)$, which is impossible.

There are exactly three types of quadruples of the sphere having only distances $r \cdot \cos^{-1}(\pm 1/3)$. They are (1) the equilateral quadruple with each distance $r \cdot \cos^{-1}(-1/3)$, (2) the center and inscribed equilateral triple of a cap of radius $r \cdot \cos^{-1}(1/3)$, and (3) the square with side $r \cdot \cos^{-1}(1/3)$ and diagonal $r \cdot \cos^{-1}(-1/3)$. Quadruples (2) and (3) are both imbeddable in a cap of radius $r \cdot \cos^{-1}(1/3)$ while quadruple (1) is imbeddable in no cap. There is a single type of pseudo-spherical quintuple and a single type of sextuple with quadruples of types (2) and (3) only, while every pseudo-spherical septuple without diametral point-pairs contains a quadruple of type (1). The figure illustrates the unique type of sextuple, the dark lines indicating a distance $r \cdot \cos^{-1}(-1/3)$ and the light $r \cdot \cos^{-1}(+1/3)$. Omitting any one of these points gives the quintuple of unique type mentioned above. The above remarks may all be verified by elementary considerations.

THEOREM 10. *The best congruence symbol for a cap* $\mathcal{K}$ *of radius* ρ, $\rho < r \cdot \cos^{-1}(1/3)$, *with respect to* Σ *is* $[4, 1]$.

Proof. 1. $\mathcal{K}$ *has the symbol* $[4, 1]$ *with respect to* Σ. Suppose that N is a semimetric space containing more than five points and that each four points of N are congruent to a subset of $\mathcal{K}$. By Theorem 9, N is then congruent to a subset N' of the sphere each four points of which can be covered by $\mathcal{K}$. Hence N' can be covered by $\mathcal{K}$ (Theorem 6), so that N is congruent to a subset of $\mathcal{K}$.

2. $\mathcal{K}$ *has not the symbol* $[4, 0]$ *with respect to* Σ. A pseudo-spherical quintuple can be constructed so small that each quadruple is imbeddable in $\mathcal{K}$. The reader will find the construction on pp. 77-78 of [5].

3. $\mathcal{K}$ *has no symbol with first index* 3 *with respect to* Σ. Any set of more than three points metrised equilaterally with a sufficiently small distance has each three points imbeddable in $\mathcal{K}$ but not the whole set.

Thus $[4, 1]$ is the best symbol for $\mathcal{K}$ with respect to Σ.

THEOREM 11. *The best congruence symbol for a cap* $\mathcal{K}$ *of radius* ρ, $r \cdot \cos^{-1}(1/3) \leq \rho < \pi r/2$, *is* $[4, 2]$ *with respect to* Σ.

Proof. 1. $\mathcal{K}$ *has the symbol* $[4, 2]$ *with respect to* Σ. Let N be a semimetric space containing more than six points, each four points of N imbeddable in $\mathcal{K}$. We show that N is congruent to a subset of the sphere which can be covered by $\mathcal{K}$. If N were not imbeddable in the sphere, then it would constitute a pseudo-spherical set of more than six points no two diametral, so that by Blumenthal's theorem all distances in N would be $r \cdot \cos^{-1}(\pm 1/3)$. Since the sphere has the symbol $[7, 0]$, some septuple of N must also be pseudo-spherical, and hence by the remarks on p. 269 N must contain an equilateral quadruple contrary to the hypothesis that each quadruple of N was imbeddable in $\mathcal{K}$. Thus N is congruent to a subset N' of the sphere each four points of which are coverable by $\mathcal{K}$. Hence $\mathcal{K}$ covers N' (Theorem 6), and N is congruently imbeddable in $\mathcal{K}$.

2. *$\mathcal{K}$ has not the symbol* $[4, 1]$ *with respect to* Σ. Each four points of the sextuple of the figure are imbeddable in $\mathcal{K}$ while the six are not.

3. *$\mathcal{K}$ has no symbol with first index* 3 for the reason given in part 3 of the proof of the preceding theorem.

Therefore $[4, 2]$ is the best symbol for $\mathcal{K}$ with respect to Σ.

THEOREM 12. *The cap of radius* $\rho < \pi r/2$ *has the congruence symbol* $[5, 0]$ *with respect to* Σ.

Proof. Since the sphere has the symbol $[5, 0]$ with respect to Σ and the cap has the symbol $[4, 0]$ with respect to the sphere, it follows that the cap has the symbol $[5, 0]$ with respect to Σ.

Remark. The symbols $[4, 2]$ *and* $[5, 0]$ *are a complete set for the cap of radius* ρ, $r \cdot \cos^{-1}(1/3) \leq \rho < \pi r/2$. For the best symbol $[4, 2]$ of the cap implies all its other symbols except $[5, 0]$.

This cap is, to our knowledge, the first example given of a space with a complete set containing more than one symbol. This is clearly possible only if the second index of its best symbol is greater than 1. A space whose best symbol has a second index equal to 1 is not unusual as both the Euclidean and hyperbolic n-dimensional spaces have the best symbol $[n + 2, 1]$, with respect to Σ.

THEOREM 13. *The best symbol for the hemisphere with respect to* Σ *is* $[5, 1]$.

Proof. 1. *The hemisphere has the symbol* $[5, 1]$ *with respect to* Σ. Since the sphere has the symbol $[5, 0]$ with respect to Σ and the hemisphere has the symbol $[5, 1]$ with respect to subsets of the sphere, it follows that the hemisphere has the symbol $[5, 1]$ with respect to Σ.

2. *The hemisphere has not the symbol* $[5, 0]$ *with respect to* Σ for each five points of three independent diametral point-pairs are on a hemisphere while the set is not.

3. *The hemisphere has no symbol with first index* 4. Consider the set of points of the sphere on the coordinate axes plus the great circle arc joining two not on the same axis. Each four points of the set are on a hemisphere while the set is not.

The results of this section are summarized in the following table:

Space	Complete set of symbols with respect to Σ	
(1) Sphere		[5, 0]
(2) Hemisphere		[5, 1]
(3) Cap of radius ρ, $r \cdot \cos^{-1}(1/3) \leq \rho < \pi r/2$	[4, 2]	[5, 0]
(4) Cap of radius ρ, $\rho < r \cdot \cos^{-1}(1/3)$	[4, 1]	

BIBLIOGRAPHY

1. E. Helly, "Über Mengen konvexer Körper mit gemeinschaftlichen Punkten," *Jahrbuch der Deutschen Mathematische Vereinigung*, vol. 32 (1923), pp. 175-176.

2. E. Steinitz, "Bedingt konvergente Reihen und konvexe Systeme," II, *Journal für die reine und angewandte Mathematik*, vol. 144 (1914), pp. 1-40.

3. K. Menger, "New foundation of Euclidean geometry," *American Journal of Mathematics*, vol. 53 (1931), pp. 721-745.

4. L. M. Blumenthal, "A new concept in distance geometry with applications to spherical subsets," *Bulletin of the American Mathematical Society*, vol. 47 (1941), pp. 435-443.

5. L. M. Blumenthal, *Distance Geometries*, The University of Missouri Studies, vol. 13, no. 2 (1938).

University of Missouri,
Columbia, Mo.

ON THE INFINITESIMAL GENERATORS OF INTEGRAL CONVOLUTIONS.*

By PHILIP HARTMAN and AUREL WINTNER.

The purpose of this paper is to transfer the principal results of the general theory of discrete convergent infinite convolutions of distribution functions to the corresponding continuous convolutions, which relate to the former in the same way as a one-parameter group relates to an infinite series. Such integral convolutions have been considered in certain degenerate cases by de Finetti;[1] in the general case, but under the restriction of a finite standard deviation, by Kolmogoroff;[2] and finally, without this restriction, by Lévy.[3]

While these investigations give an enumeration of the sheaf of distribution functions which are integral convolutions, an analysis of these distribution functions, from the point of view of the theory of real functions, has, apparently, not been attempted in the literature. Actually, it turns out that a considerable portion of the theory of discrete infinite convolutions, as developed by Jessen and Wintner,[4] is paralleled by corresponding theorems concerning integral convolutions.

Since a direct approach to the problems in question appears to lead to serious difficulties, the Lebesgue analysis of the most general integral convolution will be carried out by first developing the corresponding analysis of certain canonical components of these convolutions. To this end it is convenient to base the presentations on two exponential operators, which, at first glance, might seem to be artificial. Actually, these exponentials are nothing but the infinitesimal generators of the families considered by de Finetti and Kolmogoroff, respectively; while the general families considered by Lévy result by the adjunction of the infinitesimal translation as a third generator.

---

* Received November 11, 1940.

[1] B. de Finetti, "Le funzioni caratteristiche di legge istantanea," *Rendiconti della R. Accademia Nazionale dei Lincei*, Ser. 6, vol. 12 (1930), pp. 278-282.

[2] A. Kolmogoroff, "Sulla forma generale di un processo stocastico omogeneo," *ibid.*, vol. 15 (1932), pp. 805-808 and 866-869.

[3] P. Lévy, "Sur les intégrales dont les éléments sont des variables aléatoires indépendantes," *Annali della R. Scuola Normale Superiore di Pisa*, Ser. 2, vol. 3 (1934), pp. 337-366; cf. vol. 4 (1935), pp. 217-218. Cf. also the presentation in Lévy's book, *Théorie de l'addition des variables aléatoires*, Paris (1937), Chap. VII.

[4] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function" *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

PART I.

1. By the vectorial sum of two sets, A and B, of real numbers is meant the set, $A(+)B$, of those numbers x which have at least one representation of the form $a + b$, where a, b are numbers contained in the sets A, B, respectively. This addition is commutative, associative, and such that $A(+)B$ is the empty set if and only if at least one of the two sets A, B is empty. It is clear that $A(+)B$ contains the logical sum of A and B whenever the logical product of A and B contains $x = 0$. Hence, if A_0 denotes the logical sum of A and of the set consisting of $x = 0$ alone, and if A_n is defined by $A_n = A_{n-1} (+) A_0$, then A_n contains A_{n-1}. By A_∞ will be meant the least set containing every A_n.

It is clear that A is enumerable (at most) if and only if the same is true of A_∞. Of course, it is possible that the closure of A_∞ is the entire x-axis (this is, for instance, the case if A consists of the numbers $\pm a, \pm a^2, \pm a^3, \cdots$, where a is any fixed transcendental number greater than 1). Furthermore, A_∞ cannot be a bounded set unless A_0 consists of $x = 0$ alone. It is also clear that if A contains an interval and if the closure of the latter contains $x = 0$, then A_∞ is either the entire x-axis or one of the half-lines ending at $x = 0$.

1 bis. If $a_1\, a_2, \cdots$ are positive numbers which tend to zero and need not be distinct, let $A\{a_k\}$ denote the set of those real numbers which may be represented, in at least one way, by a convergent series of the form $\pm a_1 \pm a_2 \pm \cdots$, where the signs are chosen independently (for instance, $A\{3^{-k}\}$ is Cantor's ternary set).

Suppose that $a_1 + a_2 + \cdots < \infty$, and let b_j denote the j-th remainder, $b_j = a_{j+1} + a_{j+2} + \cdots$. It is easy to see[5] that if $a_j > b_j$ for every j, then $A\{a_k\}$ is nowhere dense, perfect, and such that

$$\text{meas } A\{a_k\} = \lim_{j\to\infty} 2^{j+1} b_j;$$

while if $a_j \leqq b_j$ for every sufficiently large j, then $A\{a_k\}$ must contain an interval.

Hence, if $0 < c < \frac{1}{2}$, then $A\{c^k\}$ is a perfect set of measure 0; while it is seen from § 1 that there exists for every c, $0 < c < 1$, an $n = n(c)$ such that $A_n\{c^k\}$ is an interval placed symmetrically about $x = 0$, which implies that $A_\infty\{c^k\}$ is the entire x-axis. However, on replacing $\{c^k\}$ by a sequence which tends sufficiently rapidly to 0, one can assure the existence of sequences $\{a_k\}$ for which not only $A\{a_k\}$ but also $A_\infty\{a_k\}$ is perfect and of measure 0.

---

[5] Cf. R. Kershner and A. Wintner, "On symmetric Bernoulli convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 541-548.

2. Let $\mu(x)$, $-\infty < x < \infty$, be a non-decreasing bounded function for which

$$\mu(-\infty) = 0; \qquad \text{so that } \mu(+\infty) = [\mu] < \infty, \qquad \text{where } [\mu] = \int_{-\infty}^{\infty} d\mu(x).$$

By the point spectrum and the spectrum of μ will be meant the sets, $P = P(\mu)$ and $S = S(\mu)$, of those real numbers, p and s, for which

$$\mu(p-0) \neq \mu(p+0) \text{ and } \mu(s-r) \neq \mu(s+r),$$

respectively, where $r(\neq 0)$ is arbitrarily small. Thus $S(\mu)$ contains $P(\mu)$ and is closed (possibly empty). The set $P(\mu)$, which is at most enumerable, is empty if and only if μ is continuous, that is, $D(\mu) = 0$, where $D(\mu)$ denotes the non-negative number

$$(1) \qquad D(\mu) = \text{maximum of } \mu(x+0) - \mu(x-0) \text{ for } -\infty < x < \infty.$$

If ν satisfies the same conditions as μ, the same is true of the convolution $\mu * \nu$, which is defined by

$$(2) \qquad \mu(x) * \nu(x) = \int_{-\infty}^{\infty} \mu(x-y)\, d\nu(y).$$

It is also clear that, with the above definition of the bracket symbol,

$$(3) \qquad [\mu * \nu] = [\mu][\nu].$$

Furthermore, it is known that, in the notations of § 1,

$$(4) \qquad P(\mu * \nu) = P(\mu)\ (+)\ P(\nu)$$

and

$$(5) \qquad S(\mu * \nu) = \text{closure of } S(\mu)\ (+)\ S(\nu).$$

Finally, it is clear from (2) that if $\{(x'_j, x''_j)\}$ is any sequence of mutually disjoint x-intervals for which

$$(6) \qquad \sum_j |x'_j - x''_j| \leqq \delta,$$

and if μ is such that (6) alone implies

$$(7) \qquad \sum_j |\mu(x'_j) - \mu(x''_j)| \leqq \epsilon,$$

then (6) also implies

$$(8) \qquad \sum_j |\mu(x'_j) * \nu(x'_j) - \mu(x''_j) * \nu(x''_j)| \leqq \epsilon[\nu].$$

3. For every real number c, let χ_c denote the particular μ which is defined by $[\chi_c] = 1$ and

(9) $S(\chi_c) = c$; so that $\chi_c(x) = \frac{1}{2} + \frac{1}{2}\operatorname{sgn}(x - c)$ and $\chi_{a+b} = \chi_a * \chi_b$,

by (2). It is also clear from (2) that, for every μ,

$$\mu(x) * \chi_c(x) = \mu(x - c). \tag{10}$$

Starting with any μ, define a sequence of μ's by the recursion formula

$$\mu^n = \mu^{n-1} * \mu, \text{ where } \mu^0 = \chi_0; \qquad (\mu^1 = \mu), \tag{11}$$

and then introduce still another μ, which we denote by exp μ and define by

$$\exp \mu = \sum_{n=0}^{\infty} \frac{\mu^n}{n!\, e^{[\mu]}}; \qquad \left([\mu] = \int_{-\infty}^{\infty} d\mu(x) \geqq 0\right). \tag{12}$$

Since it is clear from (3) and (11) that

$$0 = \mu^n(-\infty) \leqq \mu^n(x) \leqq \mu^n(\infty) = [\mu^n] = [\mu]^n, \tag{13}$$

the series (12) is uniformly convergent for $-\infty < x < \infty$; furthermore,

$$[\exp \mu] = 1. \tag{14}$$

The exponential notation is justified by the identity

$$\exp (\mu + \nu) = (\exp \mu) * (\exp \nu), \tag{15}$$

which is easily verified from the fact that the multiplication defined by (2) is commutative, associative, distributive with respect to ordinary addition, and also with respect to ordinary multiplication by (non-negative) constants.

A μ is called a distribution function if $[\mu] = 1$. While it is not assumed that this condition is satisfied, (14) shows that exp μ must be a distribution function.

4. Let $M_l(\mu)$ denote the l-th momentum,

$$M_l(\mu) = \int_{-\infty}^{\infty} x^l d\mu(x), \text{ if } \int_{-\infty}^{\infty} |x|^l d\mu(x) < \infty. \tag{16}$$

It is clear from (2) and (11) that if $M_1(\mu)$ exists, then $M_1(\mu^n)$ exists and

$$M_1(\mu^n) = nM_1(\mu)[\mu]^{n-1}, \tag{17}$$

where $M_1(\mu^0) = 0$. Similarly, if $M_2(\mu) < \infty$, then

$$M_2(\mu^n) = nM_2(\mu)[\mu]^{n-1} + n(n-1)M_1(\mu)^2[\mu]^{n-2}, \tag{18}$$

where $M_2(\mu^n) = 0$ if either $n = 0$ or $[\mu] = 0$.

Using the Stieltjes analogue to Lebesgue's theorem on term-by-term integration of monotone sequences, one easily verifies from (17), (18) and (12), that if $M_1(\mu)$ exists, then $M_1(\exp\,\mu)$ exists and

$$M_1(\exp\,\mu) = M_1(\mu); \tag{19}$$

and that if $M_2(\mu) < \infty$, then

$$M_2(\exp\,\mu) = M_2(\mu) + M_1(\mu)^2. \tag{20}$$

However, it will not be assumed that the momentum $M_1(\mu)$ or $M_2(\mu)$ of μ exists.

5. Let $P_n(\mu)$ and $S_n(\mu)$ denote the sets which, respectively, result by applying to the sets $A = P(\mu)$ and $A = S(\mu)$ of § 2 the definition of A_n in § 1, where $0 \leqq n \leqq \infty$. It is clear from (4), (5) and (11) that, if $n < \infty$, then $P_n(\mu)$ and the closure of $S_n(\mu)$, respectively, represent the point spectrum and the spectrum of

$$\sum_{k=0}^{n} p_k \mu^k,$$

whenever $p_0, \cdots, p_n$ are positive constants. It follows, therefore, from the uniform convergence of the series (12) of non-decreasing functions, that

$$P(\exp\,\mu) = P_\infty(\mu) \tag{21}$$

and

$$S(\exp\,\mu) = \text{closure of } S_\infty(\mu). \tag{22}$$

It is seen from § 1 how restricted is the class of those sets which may be represented, by a suitable A, in the form A_∞. Hence, (22) implies that only sets of rather particular structure can be spectra of distribution functions of the form $\exp\,\mu$ (notice that $A = S(\mu)$ itself must be a closed set).

6. In order to go beyond the contents of (21) and (22), it will be necessary to consider the Lebesgue analysis of μ,

$$\mu = \mu_d + \mu_a + \mu_s \equiv \mu_d + \mu_c, \tag{23}$$

where μ_d attains its total variation on an enumerable set, μ_a is absolutely continuous, while the singular component, μ_s, is continuous and attains its total variation on a set of measure 0. It will be assumed that all three components are normalized so as to tend to 0 as $x \to -\infty$. If a component does not vanish identically, μ will be said to have this component. If μ does not have more than one of its components μ_d, μ_a, μ_s, it is called a pure μ.

It is easily seen from (4) and (6)-(8) that, if ν is arbitrary and b denotes a subscript occurring in (23), then

(24) $$\mu \equiv \mu_b \text{ implies } \mu * \nu \equiv (\mu * \nu)_b \text{ for } b = c, a,$$
while
(25) $$\mu_b \not\equiv 0 \text{ implies } (\mu * \nu)_b \not\equiv 0 \text{ for } b = c, a, \qquad \text{if } \nu \not\equiv 0.$$

Since A_∞ always contains the point $x = 0$ (cf. § 1), it is indicated by (21) that sharp results can be obtained only by replacing μ by the function $\mu^\dagger$ which results by removing the possible saltus of μ at $x = 0$; so that

(26) $$\mu^\dagger = \mu - m\chi_0, \qquad \text{[cf. (9)]},$$
where
(27) $$m = \mu(+0) - \mu(-0) \geqq 0.$$
Obviously,
(28) $$(\mu + \nu)^\dagger = \mu^\dagger + \nu^\dagger.$$

Furthermore, from (15) and (26),

(29) $$\exp \mu = \exp (\mu^\dagger),$$

since (12) and (9) imply that

(30) $$\exp (c\chi_0) \equiv \chi_0.$$

6. The first two of the following three statements are clear from (21):

(i) exp μ *is always discontinuous at* $x = 0$.

(ii) (exp $\mu)^\dagger$ *is continuous if and only if the same is true of* $\mu^\dagger$.

(iii) (exp $\mu)^\dagger$ *is pure only if the same is true of* $\mu^\dagger$; *while if* $\mu^\dagger$ *is pure, the same is true of* (exp $\mu)^\dagger$ *unless* $\mu^\dagger$ *is singular.*

As to (iii), it is easily seen from (25) and (15) that it is sufficient to prove the following three statements:

(I) exp μ *is purely discontinuous if and only if the same is true of* μ (*unless* $\mu \equiv 0$, *in which case* exp $\mu = \chi_0$).

(II) (exp $\mu)^\dagger$ *is absolutely continuous if and only if the same is true of* $\mu^\dagger$.

(III) exp μ *has a singular component if and only if the same is true of* μ.

It will be shown in § 6 bis by two examples that more than (III) cannot be true, since

(III bis) exp μ *may or may not have an absolutely continuous component when* μ ($\not\equiv 0$) *is purely singular.*

Proof of (I). It is clear from (24) and (25), that $(\mu * \nu)_c \equiv 0$ if and only if $\mu_c \equiv 0$ and $\nu_c \equiv 0$. Accordingly, (11) implies that $(\mu^n)_c \equiv 0$ holds for every $n \geqq 0$ or for no $n \geqq 1$ according as it does or does not hold for $n = 1$. Hence, (I) follows from the fact that the series (12) of non-decreasing functions has positive coefficients (and is uniformly convergent for $-\infty < x < \infty$).

Proof of (II). Suppose first that μ is absolutely continuous. Then there exists, for every $\epsilon > 0$, a $\delta > 0$ for which (6) implies (7). Hence, (6) implies (8) for every ν. If $\nu = \mu^{n-1}$, then $[\nu] = [\mu]^{n-1}$, by (11) and (3). Thus

$$\sum_j \left| \sum_{n=1}^{\infty} \frac{\mu^n(x_j')}{n\,!} - \sum_{n=1}^{\infty} \frac{\mu^n(x_j'')}{n\,!} \right| \leqq \epsilon \sum_{n=1}^{\infty} \frac{[\mu]^{n-1}}{n\,!}.$$

This means that, on omitting that term of the series (12) which belongs to $n = 0$, one obtains an absolutely continuous function. It follows, therefore, from (9) that $(\exp \mu)^\dagger$ is absolutely continuous. Hence, the first statement of (II) follows from (30). The second statement of (II) is clear from the positiveness of the coefficients of the uniformly convergent series (12), since every μ^n is non-decreasing and $\mu^1 = \mu$.

Proof of (III). According to (I), (II), the functions $\exp \mu_d$, $(\exp \mu_a)^\dagger$ are purely discontinuous and absolutely continuous, respectively. Hence, it is clear from (24) and (25) that $(\exp \mu_d) * (\exp \mu_a)$ cannot have a singular component. It follows, therefore, from (15) and (23) that, if μ has no singular component, the same is true of $\exp \mu$. The converse is clear from the properties of the series mentioned at the end of the proof of (I) or (II).

6 bis. The assertion of (III bis) is two-fold, since it states that, for a suitable purely singular $\mu (\not\equiv 0)$, the function $(\exp \mu)^\dagger$, which is continuous by (ii) and is $\not\equiv 0$ by (III), may be

(III$_1$) *purely singular;*

(III$_2$) *not purely singular.*

Proof of (III$_1$). It is known [6] that if μ denotes the distribution function of the Rademacher series

$$\sum_{k=1}^{\infty} a_k \operatorname{sgn} \sin (2^k \pi t),\ 0 \leqq t \leqq 1, \text{ where } \sum_{k=1}^{\infty} a_k^2 < \infty,\ a_k > 0,$$

then μ is continuous and $S(\mu)$ is the set $A\{a_k\}$, defined in § 1 bis. It follows, therefore, from (22) and from the last remark of § 1 bis, that $\{a_k\}$ may so be chosen that not only $S(\mu)$ but also $S(\exp\ \mu)$ becomes perfect and of measure 0. But then it is clear from (ii), § 6, that not only μ but also $(\exp\ \mu)^{\dagger}$ is purely singular $(\not\equiv 0)$.

---

[6] Cf. B. Jessen and A. Wintner, *loc. cit.*[4], Theorem 11. The continuity is implied by a general theorem of P. Lévy (*loc. cit.*[13]).

Proof of (III_2). It is clear from the proof of (II), § 6, that

$$\sum_{n=k}^{\infty} \frac{\mu^n}{n!}$$

is absolutely continuous whenever the same is true of μ^k, where k is arbitrarily fixed. It follows, therefore, from (12), that if $\mu^2 \equiv \mu * \mu$ is absolutely continuous, then

$$\exp\ \mu = \text{const.}(\chi_0 + \mu + \nu), \qquad (\text{const.} = e^{-[\mu]}),$$

where ν is absolutely continuous. Hence, (III_2) is certainly true if there exists a purely singular $\mu(\not\equiv 0)$ for which $\mu * \mu$ is absolutely continuous. And the existence of such a μ can be assured as follows:

Let $L(u;\mu)$, $-\infty < u < \infty$, denote the Fourier-Stieltjes transform of μ; so that, by Fubini's theorem,

$$L(u;\mu * \nu) = L(u;\mu)L(u;\nu).$$

It is known[7] that there exists for *every* $\alpha < \frac{1}{2}$ a purely singular $\mu(\not\equiv 0)$ for which

$$L(u;\mu) = O(|\,u\,|^{-\alpha}) \qquad \text{as } u \rightarrow \pm\,\infty.$$

Hence, if $\alpha = \frac{1}{3}$, then

$$L(u;\mu * \mu) = O(|\,u\,|^{-2/3}), \text{ and so} \int_{-\infty}^{\infty} |\,L(u;\mu * \mu)\,|^2\, du < \infty.$$

It follows, therefore from Plancherel's theorem that $\mu * \mu$ is absolutely continuous.[8]

---

[7] N. Wiener and A. Wintner, "Fourier-Stieltjes transforms and singular infinite convolutions," *American Journal of Mathematics*, vol. 60 (1938), pp. 513-522. While only Fourier-Stieltjes momenta, instead of Fourier-Stieltjes transforms, are considered there, the construction applies automatically to the latter case also.

[8] With a derivative of class $L_2(-\infty, +\infty)$. Similarly, if $\alpha \rightarrow \frac{1}{2} - 0$, it follows from Titchmarsh's extension of Plancherel's theorem in the Hölder-Hausdorff direction that, for every fixed $p \geqq 2$, the derivative of $\mu * \mu$ can be of class $L_r(-\infty, +\infty)$ for every r between $r = 2$ and $r = p$. Incidentally, it may be shown that there exists a fixed μ for $\alpha = \frac{1}{2}$; so that the derivative of a fixed $\mu * \mu$ can be of class $L_p(-\infty, +\infty)$ for every $p \geqq 2$.

PART II.

7. Starting with an arbitrary μ, remove its saltus at $x = 0$ (if any). Decompose the resulting function, (26), by placing[9]

$$\mu^{\dagger} = \sum_{j=1}^{\infty} \mu'_j, \tag{31}$$

where every μ'_j is a μ and does not contain $x = 0$ in its spectrum. It is clear that such a decomposition of $\mu^{\dagger}$ is always possible, and that the series (31) of non-negative, non-decreasing, uniformly bounded functions μ'_j is uniformly convergent for $-\infty < x < \infty$, no matter how the decomposition is chosen.

Since $x = 0$ is not in the spectrum of μ'_j,

$$\mu''_j(x) = \int_{-\infty}^{x} y^{-2} d\mu'_j(y), \qquad -\infty < x < \infty, \tag{32}$$

defines[10] a function μ''_j which is again a μ, does not contain $x = 0$ in its spectrum, and possesses the finite second momentum

$$M_2(\mu''_j) = \int_{-\infty}^{\infty} d\mu'_j(x) = [\mu'_j], \tag{33}$$

where M_l, [ ] are the operators defined in § 4, § 2, respectively.

Let c_j denote the first momentum of μ''_j; so that

$$M_1(\mu''_j) = \int_{-\infty}^{\infty} x^{-1} d\mu'_j(x) = c_j, \tag{34}$$

by (32). According to (33) and (20),

$$M_2(\exp \mu''_j) = [\mu'_j] + c_j^2, \tag{35}$$

since, by (34) and (19),

$$M_1(\exp \mu''_j) = c_j. \tag{36}$$

Finally, introduce, for every term μ'_j of (31), a distribution function, $\bar{\mu}_j$, which is defined by means of (32) and (34), as follows:

$$\bar{\mu}_j = \chi_{-c_j} * \exp \mu''_j. \tag{37}$$

---

[9] The primes do not, of course, refer to differentiations.

[10] Cf. the preceding footnote.

It is clear from (10) that $\bar{\mu}_j$ differs from the distribution function $\exp \mu''_j$ only in that the origin is transferred to the center of mass (36). Thus

$$M_1(\bar{\mu}_j) = 0. \tag{38}$$

Consequently, (35) goes over into

$$M_2(\bar{\mu}_j) = [\mu'_j], \tag{39}$$

since a second momentum attains its minimum when the center of mass is the origin.

8. Since it is clear from (31) that

$$\sum_{j=1}^{\infty} [\mu'_j] = [\mu^\dagger], \text{ where } 0 \leqq [\mu^\dagger] < \infty, \tag{40}$$

it follows from (38) and (39) that the convolution

$$\bar{\mu} = \bar{\mu}_1 * \bar{\mu}_2 * \cdots * \bar{\mu}_j * \cdots \tag{41}$$

is absolutely convergent.[11]

The notation (41) is justified by the fact that the distribution function $\bar{\mu}$ depends only on the given μ (actually, only on $\mu^\dagger$), and *not* on the particular choice of the terms μ'_j in the decomposition (31) of the function which is defined by (26) and (27). In fact, if

$$\mu^\dagger = \sum_{j=1}^{\infty} \mu`_j \tag{42}$$

is another decomposition, one can obviously unite (31) and (42) into one decomposition of $\mu^\dagger$. Since the infinite convolution belonging to this united decomposition is absolutely convergent, it may be rearranged[12] into the absolutely convergent infinite convolution which belongs to either of the decompositions (31), (42). This clearly implies that $\bar{\mu}$ is uniquely determined by $\mu^\dagger$.

It is also seen from this proof that if ν is another μ, then

$$\bar{\rho} = \bar{\mu} * \bar{\nu}, \text{ where } \rho = \mu + \nu. \tag{43}$$

This suggests the introduction of the notation $\bar{\mu} = \text{Exp } \mu^\dagger$; so that

$$\text{Exp } \mu^\dagger = \bar{\mu}_1 * \bar{\mu}_2 * \cdots * \bar{\mu}_j * \cdots; \qquad \text{cf. (41).} \tag{44}$$

In fact, (43) then appears in the form

$$\text{Exp } (\mu^\dagger + \nu^\dagger) = (\text{Exp } \mu^\dagger) * (\text{Exp } \nu^\dagger); \qquad \text{cf. (28).} \tag{45}$$

---

[11] Cf. B. Jessen and A. Wintner, *loc. cit.*[4], Theorems 4 and 6.

[12] Cf. B. Jessen and A. Wintner, *loc. cit.*[4], p. 58.

Since (44) is a convergent (and even an absolutely convergent) convolution of distribution functions, it is a distribution function; so that

$$(46) \qquad [\mathrm{Exp}\ \mu^{\dagger}] = 1.$$

Needless to say, Exp $\mu^{\dagger}$ means Exp (μ), where $\mu = \mu^{\dagger}$, and *not* (Exp $\mu)^{\dagger}$, where $\mu = \mu^{\dagger}$ (these two functionals of $\mu^{\dagger}$ are not identical, since Exp $(\mu^{\dagger})$ can have a saltus at $x = 0$). Accordingly, it would be possible to denote Exp $\mu^{\dagger} \equiv$ Exp $(\mu^{\dagger})$ simply by Exp μ; cf. (28). This simplified notation will not be used for the reason that, while the function Exp μ is now defined only if $\mu = \mu^{\dagger}$, it will later (§ 10) be defined for $\mu \neq \mu^{\dagger}$ also; and this extension of the domain of Exp μ will be such that

$$(47) \qquad \mathrm{Exp}\ \mu \neq \mathrm{Exp}\ (\mu^{\dagger}) \neq (\mathrm{Exp}\ \mu)^{\dagger} = \mathrm{Exp}\ \mu, \text{ if } \mu \neq \mu^{\dagger}.$$

9. In order to determine the spectrum of Exp $\mu^{\dagger}$, use will be made of the known fact[13] that if $\sigma_1 * \sigma_2 * \cdots$ is a convergent convolution of distribution functions, then

$$(48) \qquad S(\sigma_1 * \sigma_2 * \cdots) = S(\sigma_1)\ (+)\ S(\sigma_2)\ (+) \cdots,$$

where the infinite vectorial sum is defined as the set of those values s for which there exists at least one representation of the form

$$(48 \text{ bis}) \qquad s = \lim_{j \to \infty} s_j, \text{ where } s_j \text{ is in } S(\sigma_1)\ (+) \cdots (+)\ S(\sigma_j).$$

Since it is clear from (32) that $S(\mu''_j) = S(\mu'_j)$, it follows from (22) and (37) that

$$(49) \qquad S(\bar{\mu}_j) = S(\chi_{-c_j})\,(+) \text{ closure of } S_{\infty}(\mu'_j).$$

Thus, from (44) and (48),

$$(50) \qquad S(\mathrm{Exp}\ \mu^{\dagger}) = [S(\chi_{-c_1})\,(+) \text{ closure of } S_{\infty}(\mu'_1)]\,(+) \\ [S(\chi_{-c_2})\ (+) \text{ closure of } S_{\infty}(\mu'_2)]\,(+) \cdots.$$

10. For every constant $s \geqq 0$, let κ_s denote the symmetric normal distribution of standard deviation $\sqrt{s}$; so that, by a known (and characteristic) property of these distribution functions,

$$(51) \qquad \kappa_r * \kappa_s = \kappa_{r+s}.$$

It is understood that κ_0 is defined to be χ_0; cf. (9). Thus

$$(52) \qquad S(\kappa_m) = (-\infty, \infty) \text{ if } m > 0, \text{ while } S(\kappa_0) = 0.$$

---

[13] B. Jessen and A. Wintner, *loc. cit.*[4], Theorem 3.

In § 8, Exp μ was defined only for the case $\mu = \mu^\dagger$; cf. (26), (27). For the general case, define Exp μ by

(53) $$\operatorname{Exp} \mu = \kappa_m * \operatorname{Exp} \mu^\dagger, \text{ where } m = \mu(+0) - \mu(-0) \geqq 0.$$

This definition is consistent, since $\kappa_0 = \chi_0$. Furthermore,

(54) $$\operatorname{Exp}\ (m\chi_0) = \kappa_m.$$

Finally, the functional equation is preserved; in fact,

(55) $$\operatorname{Exp}\ (\mu + \nu) = (\operatorname{Exp}\ \mu) * (\operatorname{Exp}\ \nu)$$

follows from (28), (45), (51) and (53).

The spectrum of (53) is supplied by (50) and (52), since

(56) $$S(\operatorname{Exp}\ \mu) = S(\kappa_m)(+)S(\operatorname{Exp}\ \mu^\dagger),$$

by (5). Similarly, from (4) and (53),

(57) $$P(\operatorname{Exp}\ \mu) = \text{empty set unless } \mu = \mu^\dagger.$$

Finally, from (3), (46) and (53),

(58) $$[\operatorname{Exp}\ \mu] = 1.$$

11. Let $L(u;\nu)$ denote the Fourier-Stieltjes transform

(59) $$L(u;\nu) = \int_{-\infty}^{\infty} e^{iux} d\nu(x), \qquad -\infty < u < \infty; \qquad (L(0;\nu) = [\nu]).$$

It is known that, in view of Fubini's theorem, (3) may be generalized as follows:

(60) $$L(u;\mu * \nu) = L(u;\mu)L(u;\nu); \qquad \text{cf. } (2).$$

In particular

(61) $$L(u;\chi_a * \nu) = e^{iau}L(u;\nu),$$

since, by (9),

(62) $$L(u;\chi_a) = e^{iau}.$$

It will now be verified that if $\nu = \exp \mu$, then the continuous function (59) cannot vanish, and that the branch of its logarithm which is real ($= 0$) at $u = 0$ is

(63) $$\log L(u;\ \exp\ \mu) = \int_{-\infty}^{\infty} (e^{iux} - 1)\, d\mu(x).$$

Similarly, the logarithm of the Fourier-Stieltjes transform of Exp μ turns out to be[14]

---

[14] Cf. A. Kolmogoroff, *loc. cit.*[2]

$$(64) \qquad \log L(u;\ \operatorname{Exp}\mu) = \int_{-\infty}^{\infty} \frac{e^{iux}-1-iux}{x^2}\, d\mu(x),$$

a transform of the type on which Wiener's generalized harmonic analysis is based.[15] It is understood that the value of the kernel at $x = 0$ is meant to be

$$(64\text{ bis}) \qquad \lim_{x\to\pm 0} \frac{e^{iux}-1-iux}{x^2} \equiv -\tfrac{1}{2}u^2 \qquad (\text{also when } u = 0).$$

First, it is clear from (60), (62) and (11) that $L(u;\mu^n) = (L(u;\mu))^n$. Hence, (63) follows from (59) and (12) by a trivial term-by-term integration.

Next, from (63) and (32),

$$\log L(u;\exp \mu''_j) = \int_{-\infty}^{\infty} \frac{e^{iux}-1}{x^2}\, d\mu'_j(x) \equiv ic_j u + \int_{-\infty}^{\infty} \frac{e^{iux}-1-iux}{x^2}\, d\mu'_j(x),$$

by (34). Hence, from (37) and (61),

$$(66) \qquad \log L(u;\ \bar\mu_j) = \int_{-\infty}^{\infty} \frac{e^{iux}-1-iux}{x^2}\, d\mu'_j(x).$$

Consequently, from (60),

$$\log L(u;\ \bar\mu_1 * \cdots * \bar\mu_n) = \int_{|x|\geqq 1/n} \frac{e^{iux}-1-iux}{x^2}\, d\mu(x),$$

if the μ'_j of the decomposition (31) are so chosen that $d\mu'_j(x) \equiv d\mu(x)$ or $d\mu'_j(x) \equiv 0$ according as $j^{-1} \leqq |x| < (j-1)^{-1}$ does or does not hold, where $j = 1, 2, \cdots$ and $0^{-1} = +\infty$. Accordingly, from (44),

$$\log L(u;\ \operatorname{Exp}\mu^\dagger) = \lim_{n\to\infty} \int_{|x|\geqq 1/n} \frac{e^{iux}-1-iux}{x^2}\, d\mu(x),$$

by Lévy's continuity theorem of the Fourier-Stieltjes transforms of distribution functions.

The last relation shows that (64) is true for $\mu = \mu^\dagger$. Since the Fourier-Stieltjes transform of the normal distribution function (54) is known to be $e^{-\frac{1}{2}mu^2}$, it is clear from (64 bis) that (64) is true for $\mu = m\kappa_0$ also. It follows, therefore, from (53) and (60) that (64) is true for every μ.

It should be noted for later application that the real parts of (63) and (64) can be represented, respectively, by

---

[15] Cf., e. g., N. Wiener, *The Fourier integral and certain of its applications*, Cambridge, 1933, Chap. IV.

$$-\tfrac{1}{2}\log|L(2u;\,\exp\,\mu)| = \int_{-\infty}^{\infty}(\sin\,ux)^2 d\mu(x) \tag{67}$$

and

$$-\tfrac{1}{2}\log|L(2u;\,\mathrm{Exp}\,\mu)| = \int_{-\infty}^{\infty}\left(\frac{\sin\,ux}{x}\right)^2 d\mu(x). \tag{68}$$

12. The relations (50), (54), (55), (58) correspond to (22), (30), (15), (14), respectively. On the other hand, the italicized statements of § 6 do not remain valid in any sense, if exp is replaced by Exp. In fact, it will now be shown by a straightforward adaptation of a method previously used[16] in case of convergent convolutions $\nu = \sigma_1 * \sigma_2 * \cdots$ of distribution functions, that the distribution function $\nu = \mathrm{Exp}\,\mu^{\dagger}$ can possess an arbitrarily high degree of smoothness, whether $\mu^{\dagger}$ be pure of any of the three types or mixed.

(I). *If there exists a positive constant* $\lambda < 2$ *for which*

$$\int_{-\infty}^{\infty} |x|^{-\lambda} d\mu(x) = \infty, \tag{69}$$

then Exp μ *has, for* $-\infty < x < \infty$, *derivatives of arbitrarily high order, which all tend to* 0 *as* $x \to \pm\infty$. *If* $\lambda < 1$, *then* Exp μ *is a transcendental entire function.*

It is understood that $\mu \not\equiv \mu^{\dagger}$ is sufficient but not necessary for (69); and that (69) is equivalent to

$$\int_{-\epsilon}^{\epsilon} |x|^{-\lambda} d\mu(x) = \infty, \qquad \epsilon > 0, \qquad (0 < \lambda < 2),$$

since μ is a bounded monotone function. Accordingly, what is stated to impose on Exp μ a high degree of smoothness is nothing but a sufficient amount of steepness of the curve $\mu = \mu(x)$ near $x = 0$. It will be shown in the proof of the Corollary to (II), § 13 below, that if the question of the analyticity of Exp μ is replaced by the question of its mere continuity, then one is led to a similar, though of course less pronounced, assumption concerning the steepness of $\mu = \mu(x)$ near $x = 0$; in fact, it will then be sufficient to require the divergence of the integral (69) exactly in the limiting case $\lambda = 2$. Cf. (III), § 14 below.

---

[16] A. Wintner, "On symmetric Bernoulli convolutions," *Bulletin of the American Mathematical Society*, vol. 41 (1935), 137-138; "On the smoothness of infinite convolutions of the type occurring in the theory of the Riemann zeta function," *American Journal of Mathematics*, vol. 61 (1939), pp. 231-236; cf. P. Erdős, "On the smoothness of the asymptotic distribution of additive arithmetical functions," *ibid.*, pp. 722-726.

Since the integrand of (68) is non-negative,

$$(70)\quad \liminf_{u\to\pm\infty} \frac{-1}{2\,|u|^{2-\lambda-\epsilon}} \log|L(2u;\operatorname{Exp}\mu)| \geqq \liminf_{u\to\pm\infty} \frac{1}{|u|^{2-\lambda-\epsilon}} \int_{-1/|u|}^{1/|u|} \left(\frac{\sin ux}{x}\right)^2 d\mu(x),$$

where ϵ is arbitrary. But it is clear that there exists a positive constant such that

$$(71)\qquad |\sin y| \geqq \text{const.}\, |y| \quad \text{for } -1 \leqq y \leqq 1.$$

Hence, on placing $y = ux$, one sees that the expression on the right of (70) is not less than a constant multiple of

$$\liminf_{u\to\pm\infty} \frac{1}{|u|^{2-\lambda-\epsilon}} \int_{-1/|u|}^{1/|u|} \frac{(ux)^2}{x^2}\, d\mu(x) = \liminf_{u\to\pm\infty} |u|^{\lambda+\epsilon}\{\mu(1/|u|) - \mu(-1/|u|)\}.$$

Since the assumption (69) obviously implies that

$$\lim_{u\to\pm\infty} |u|^{\lambda+\epsilon}\{\mu(1/|u|) - \mu(-1/|u|)\} = \infty$$

for every fixed $\epsilon > 0$, it follows that the expression on the left of (70) is ∞. This means that

$$(72)\qquad L(u;\operatorname{Exp}\mu) = O(e^{-|u|^{2-\lambda-\epsilon}}), \qquad u \to \pm\infty,$$

for every fixed $\epsilon > 0$. Hence the statements of the italicized theorem follow from known connections between the smoothness of ν and an exponential estimate of (59) as $u \to \pm\infty$.

13. The absolute continuity of Exp μ is not likely to be characterizable in terms of explicit properties of μ. The following pair of conditions involve only an averaged order of μ. The difficulties of the gap between these two conditions appear to be of diophantine irregularity and are rather obscure.

(II). *In order that* Exp μ *be absolutely continuous, it is necessary that*

$$(73)\qquad \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} \left(\frac{\sin x}{\epsilon x}\right)^2 d\mu(\epsilon x) = \infty$$

and it is sufficient that

$$(74)\qquad \liminf_{\epsilon\to 0} \frac{1}{|\log \epsilon|} \int_{-\infty}^{\infty} \left(\frac{\sin x}{\epsilon x}\right)^2 d\mu(\epsilon x) > \tfrac{1}{2}$$

(the integrations refer to x). In fact, if l is any number less than the expression (74), *then* Exp μ *has at least* $[2l]$ *continuous derivatives which tend to* 0 *as* $x \to \pm\infty$.

Corollary. *If*

$$\int_{-\infty}^{\infty} \frac{d\mu(x)}{x^2} = \infty \tag{75}$$

but[17]

$$\liminf_{\epsilon \to 0} \int_{-\infty}^{\infty} \left(\frac{\sin x}{\epsilon x}\right)^2 d\mu(\epsilon x) < \infty, \tag{76}$$

then Exp μ *is continuous but not absolutely continuous.*

First, it is clear from (68) that (73) is necessary and sufficient for

$$L(u;\nu) \to 0, \qquad u \to \pm\infty, \tag{77}$$

where $\nu = \text{Exp}\ \mu$. But (59) shows that (77) is, in view of the integral analogue to the Riemann-Lebesgue lemma, a necessary condition for the absolute continuity of ν; while (77) is a sufficient (though not a necessary) condition for the continuity of ν.

Next, suppose that the expression on the left of (74) exceeds a positive number, l. Then (68) shows that there exists a sufficiently small $\delta > 0$ such that

$$L(u;\nu) = O(|u|^{-2l-\delta}), \qquad u \to \pm\infty, \tag{78}$$

is satisfied by $\nu = \text{Exp}\ \mu$. But it is known that (78) implies for ν the degree of smoothness stated after (74).

Finally, the Corollary follows from the fact that, according to (III), § 14 below, Exp μ must be continuous if (75) is satisfied.

In order to illustrate the Corollary, let $\mu = \mu(x)$ be the step function which has at $x = \pm a^j$ the saltus a^{2j}, where $0 < a < 1$ and $j = 1, 2, \cdots$. Then (75) is obviously satisfied; while (76) requires that

$$\liminf_{\epsilon \to 0} \sum_{j=1}^{\infty} \sin^2(a^j \epsilon^{-1}) < \infty.$$

This is certainly true if a is the reciprocal value of an integer, since then

$$\sum_{j=1}^{\infty} \sin^2(a^{j-n}\pi) = \sum_{j=1}^{\infty} \sin^2(a^j\pi), \quad \text{where} \quad n = 1, 2, \cdots.$$

---

[17] N. Wiener "The spectrum of an array," *Journal of Mathematics and Physics*, vol. 6 (1927), pp. 145-157 (in particular, p. 151).

14. It will now be shown that, in contrast with the problem of absolute continuity, the question concerning the mere continuity of Exp μ can be answered completely. The result to be obtained corresponds to (ii), § 6 but lies very much deeper. In fact, the proof will depend on a fundamental theorem of Lévy,[18] which states that, for any convergent convolution $\sigma_1 * \sigma_2 \cdots$ of distribution functions σ_j,

$$D(\sigma_1 * \sigma_2 * \cdots) = 0 \text{ if and only if } \prod_{j=1}^{\infty} D(\sigma_j) = 0, \tag{79}$$

where $D(\mu)$ denotes the maximum saltus of μ; cf. (1).

(III). *The function* Exp μ *is continuous if and only if* $\int_{-\infty}^{\infty} x^{-2} d\mu(x) = \infty$.

Cf. the comments made after (I), § 12.

The statement is that

$$D(\text{Exp } \mu) = 0 \text{ if and only if } \int_{-\infty}^{\infty} x^{-2} d\mu(x) = \infty. \tag{80}$$

Suppose that (80) has already been proved for every continuous μ, say $\mu = \mu^I$, and for every purely discontinuous μ, say $\mu = \mu^{II}$. Then

$$D(\text{Exp } \mu^I) D(\text{Exp } \mu^{II}) = 0 \text{ if and only if } \int_{-\infty}^{\infty} x^{-2} d\{\mu^I(x) + \mu^{II}(x)\} = \infty.$$

It follows, therefore, from (55) that (80) is true for $\mu = \mu^I + \mu^{II}$ also, provided that it is true that

$$D(\nu^I * \nu^{II}) = 0 \text{ if and only if } D(\nu^I) D(\nu^{II}) = 0.$$

Since (1) and (4) imply that this is true, it follows that it is sufficient to prove (80) for every continuous μ and for every purely discontinuous μ.

Furthermore, (57) shows that (80) is certainly true if $\mu \neq \mu^{\dagger}$. Hence, it is sufficient to prove (80) for the case where $\mu = \mu^{\dagger}$ and $\mu^{\dagger}$ is either continuous or purely discontinuous. It will turn out that the decomposition (31) can in either case be so chosen that

$$D(\bar{\mu}_j) = e^{-q_j}, \text{ where } q_j = [\mu''_j] \equiv \int_{-\infty}^{\infty} d\mu''_j(x). \tag{80 bis}$$

It is clear from the definitions (31), (32), (44) and from Lévy's theorem

---

[18] P. Lévy, "Sur les séries dont les termes sont des variables éventuelles indépendentes," *Studia Mathematica*, vol. 3 (1931), pp. 119-155, Théorème XIII.

(79), that (80 bis) implies (80) for $\mu = \mu^\dagger$. Thus it will be sufficient to ascertain (80 bis) for the two cases mentioned.

In the first case, where $\mu^\dagger$ is continuous, each of the functions μ'_j occurring in a decomposition (31) of $\mu^\dagger$ is continuous, and so the same is true of the functions (32). It follows, therefore, from (11) and (4) that $\mu''_j{}^n$ is continuous for every $n \neq 0$. Consequently, it is clear from (1) and from the uniformity of the convergence of the series (12) for $-\infty < x < \infty$, that, if q_j denotes the number $[\mu''_j]$,

$$D(\exp \mu''_j) = D(e^{-q_j}\mu''_j{}^0) = e^{-q_j}D(\chi_0) = e^{-q_j},$$

by (9). Hence, (80 bis) follows from (37) and (10).

In the second case, where $\mu^\dagger$ is purely discontinuous, the decomposition (31) can be so chosen that, for every j,

(81) $$\mu'_j = p_j\chi_{a_j}, \text{ where } p_j \geqq 0 \text{ and } a_j \gtrless 0.$$

The possibility of this choice is clear from (9), since the bounded non-decreasing function $\mu^\dagger$ now consists of a sequence of jumps and is, by its definition (26)-(27), continuous at $x = 0$.

In addition to (81), it can be assumed that

(81 bis) $$p_j < a_j^2, \qquad (p_j \geqq 0),$$

for every j. In fact, if a $(\neq 0)$ is any number in the point spectrum of $\mu^\dagger$, one can choose an integer $N = N(p, a)$ so large that $p = \mu(a+0) - \mu(a-0)$ is less than Na^2; in which case the jump p at $x = a$ can be represented as the sum of N terms of the form (81), where $a_j = a$ for each of the N values of j, while the sum of the N coefficients p_j is p.

It is clear from (81), (32) and (9) that $\mu''_j = q_j\chi_{a_j}$, where $q_j = p_j/a_j^2$; hence, $0 \leqq q_j < 1$, by (81 bis). On the other hand, $\mu''_j = q_j\chi_{a_j}$ implies that $[\mu''_j] = q_j$, by (9); so that, from (12),

$$\exp \mu''_j = \sum_{n=0}^{\infty} \frac{(q_j\chi_{a_j})^n}{n!e^{q_j}} = \sum_{n=0}^{\infty} \frac{q_j{}^n\chi_{na_j}}{n!e^{q_j}}.$$

This means that $\exp \mu''_j$ is the step function which has at $x = na_j$ the saltus $n!^{-1}q_j{}^n e^{-q_j}$, where $n = 0, 1, 2, \cdots$. It follows, therefore, from $0 \leqq q_j < 1$ that the greatest saltus of $\exp \mu''_j$ amounts to $0!^{-1}q_j{}^0 e^{-q_j} = e^{-q_j}$ (also when $q_j = 0$). Since $q_j = [\mu''_j]$, and since (37) and (10) imply that $\bar{\mu}_j$ has the same greatest saltus as μ''_j, the proof of (80 bis) is complete.

15. In view of the vagueness of the parallelism between exp and Exp, it is somewhat unexpected that the theorem just proved can be refined as follows:

(III bis). *The function* Exp μ *is not continuous if and only if*

(82)
$$\mu''(x) = \int_{-\infty}^{x} y^{-2} d\mu(y)$$

defines a (non-decreasing) bounded function for $-\infty < x < \infty$; *in which case*

(83)
$$\text{Exp } \mu = \chi_{-c} * \exp \mu'',$$

where

(84)
$$c = \int_{-\infty}^{\infty} x^{-1} d\mu(x)$$

then exists absolutely.

COROLLARY. *If* Exp μ *is discontinuous, it must have a saltus at the particular point* $x = c$.

In fact, it is clear from (III), § 14 that (82) does not define a bounded function if and only if Exp μ is continuous. Suppose that Exp μ is not continuous. Then $\mu = \mu^{\dagger}$, by (57). Let the decomposition (31) be chosen as in § 11; so that

$$d\mu'_j(x) = \begin{cases} d\mu(x) \text{ if } j^{-1} \leqq |x| < (j-1)^{-1}, \\ 0 \text{ otherwise}; \end{cases} \qquad (j = 1, 2, \cdots; 0^{-1} = +\infty).$$

Then, by (34),

$$c_j = \int_{j^{-1} \leqq x < (j-1)^{-1}} x^{-1} d\mu(x) = M_1(\mu''_j).$$

Since the assumption, $\int_{-\infty}^{\infty} x^{-2} d\mu(x) < \infty$, implies that $\int_{-\infty}^{\infty} |x|^{-1} d\mu(x) < \infty$, it follows that the series $c_1 + c_2 + \cdots$ converges absolutely and represents the value c which is defined by (84). Hence, it is clear from (37), (44) and (10) that, since $\mu^{\dagger} = \mu$,

$$\chi_c * \text{Exp } \mu = (\exp \chi''_1) * (\exp \mu''_2) * \cdots;$$

in fact, this convolution of the distribution functions $\exp \mu''_j$ is absolutely convergent, since the same is always true of (44); cf. the statement made after (41). Now, by (60),

$$L(u; \chi_c * \text{Exp } \mu) = \prod_{j=1}^{\infty} L(u; \exp \mu''_j),$$

in virtue of Lévy's continuity theorem of the Fourier-Stieltjes transform. Thus it is clear from the formula which precedes (66), that

$$\log L(u;\chi_c * \operatorname{Exp}\mu) = \sum_{j=1}^{\infty}\int_{-\infty}^{\infty}\frac{e^{iux}-1}{x^2}\,d\mu_j(x) = \sum_{j=1}^{\infty}\int_{j^{-1}\leqq x<(j-1)^{-1}}\frac{e^{iux}-1}{x^2}\,d\mu(x),$$

by the above choice of the μ'_j. It follows, therefore, from (82), where $\mu=\mu^\dagger$, that

$$\log L(u;\chi_c * \operatorname{Exp}\mu) = \int_{-\infty}^{\infty}(e^{iux}-1)\,d\mu''(x).$$

This, when compared with (63), completes the proof of (83).

Finally, the Corollary stated after (84) is clear from (83), (10) and (i), § 6.

16. According to § 4, the distribution function exp μ has a finite second momentum if and only if $M_2(\mu)<\infty$. On the other hand, it will now be shown that the distribution function Exp μ has a finite second momentum for every μ; in fact,

(85) $$M_2(\operatorname{Exp}\mu) = [\mu],$$

while

(86) $$M_1(\operatorname{Exp}\mu) = 0,$$

no matter what μ be.

It is known[19] that if $\sigma_1 * \sigma_2 * \cdots$ is a convergent convolution of distribution functions σ_j each of which has a vanishing first momentum and a second momentum $M_2(\sigma_j)$, then the convergence of the series $M_2(\sigma_1)+M_2(\sigma_2)+\cdots$ implies that $M_2(\sigma_1 * \sigma_2 * \cdots)$ is finite and has the value represented by this series, while $M_1(\sigma_1 * \sigma_2 * \cdots)=0$. On applying this to the convergent convolution (44), one see from (38), (39) and (40), that (85) and (86) are true for $\mu=\mu^\dagger$. On the other hand, it is clear from (54) that (85) and (86) are true for the symmetric normal distribution Exp $\mu=\kappa_m$ also. Since both the squares of the standard deviations and the first momenta are additive on convolution, it follows from (53) that (85) and (86) are true for every μ.

PART III.

17. For later application, there will now be collected the distribution functions exp μ, Exp μ belonging to a μ which consists of a single saltus.

It is clear from (9)-(12) that

(87) $$\exp(q\chi_c) = \sum_{n=0}^{\infty}\frac{q^n\chi_{nc}}{n!e^q},\quad \text{where}\quad q\geqq 0,\ c\gtreqless 0$$

---

[19] B. Jessen and A. Wintner, *loc. cit.*[4], Theorem 4.

are constants. On the other hand, if the constants p, a are given and q, c are defined by

$$(88) \qquad q = \frac{p}{a^2}, \qquad c = \frac{p}{a} \quad \text{where} \quad p \geqq 0, a \gtrless 0,$$

then the function (82) and the constant (84) belonging to $\mu = p\chi_a$ are $q\chi_a$ and c, respectively. It follows, therefore, from (83) that

$$(89) \qquad \operatorname{Exp}(p\chi_a) = \chi_{-c} * \exp(q\chi_a), \text{ if } a \gtrless 0; \text{ while } \operatorname{Exp}(p\chi_0) = \kappa_p,$$

by (54).

The content of (88) and (89) is that, except for the limiting case $a = 0$ of a symmetric normal distribution function, $\operatorname{Exp}(p\chi_a)$ may be obtained from a Poisson distribution function of standard deviation $\sqrt{p/a^2}\,(\geqq 0)$ by those choices of the origin, orientation and unit of length (on the x-axis) which are assigned by (86) and (85). In fact, it is seen from (87) that, while $\exp(q\chi_0) = \chi_0$ for every q, the functions $\exp(q\chi_1)$ and $\exp(q\chi_{-1})$ represent that Poisson distribution and its conjugate, respectively, which have the standard deviation $\sqrt{q}\,(\geqq 0)$. It is understood that by the distribution function conjugate to a given distribution function $\sigma(x)$ is meant the function $1 - \sigma(-x)$.

It is seen from (63) and (64), respectively, that

$$(87 \text{ bis}) \qquad \log L(u;\ \exp(q\chi_o)) = q(e^{iou} - 1)$$

and

$$(89 \text{ bis}) \quad \log L(u;\ \operatorname{Exp}(p\chi_a)) = \begin{cases} -icu + q(e^{iau} - 1), \text{ if } a \gtrless 0; \text{ cf. (88)}, \\ -\tfrac{1}{2}pu^2, \text{ if } a = 0; \text{ cf. (64 bis)}. \end{cases}$$

18. If exp is replaced by Exp in the italicized statements of § 6, all that remains true may be summarized as follows:

As in (23), let μ_d, μ_a, μ_s, $\mu_c = \mu_a + \mu_s$, denote, respectively, the purely discontinuous, the absolutely continuous, the singular, and the continuous components of an arbitrary μ; so that, by (55),

$$(90) \qquad \begin{aligned} \operatorname{Exp}\mu &= (\operatorname{Exp}\ \mu_d) * (\operatorname{Exp}\ \mu_c), \text{ since } \mu = \mu_d + \mu_c, \\ \operatorname{Exp}\mu_c &= (\operatorname{Exp}\ \mu_a) * (\operatorname{Exp}\ \mu_s), \text{ since } \mu_c = \mu_a + \mu_s. \end{aligned}$$

Then

(I) *if* Exp μ *is purely discontinuous, then* $\mu = \mu^\dagger$ *and* $\mu_c \equiv 0$;

(II) *if* Exp μ *has no absolutely continuous component, then* $\mu = \mu^\dagger$ *and* $\mu_a \equiv 0$, *while neither* Exp μ_d *nor* Exp μ_s *has an absolutely continuous component*;

(III) *if* Exp μ *is purely singular, then either* Exp μ_d *or* Exp μ_s *is purely singular*.

Proof of (I). Suppose that Exp μ is purely discontinuous. Then $\mu = \mu^\dagger$, by (53) and the case $b = c$ of (24). Furthermore, (III bis), § 15 assures us that there exists a number c for which the function (82) satisfies (83). Since Exp μ is purely discontinuous, it is clear from (10) and (83) that the same is true of exp μ''. It follows, therefore, from (I), § 6 that μ'' is purely discontinuous. Hence, it is clear from (82) that μ is purely discontinuous.

Proof of (II). Suppose that Exp μ has no absolutely continuous component. Then $\mu = \mu^\dagger$, by (53) and the case $b = a$ of (25). Suppose, if possible, that $\mu = \mu^\dagger$ has an absolutely continuous component. Then the decomposition (31) can be so chosen that μ'_1 is absolutely continuous but not a constant; in which case it is clear from (II), § 6 that exp μ'_1 has an absolutely continuous component. Hence, (37) and (10) show that the same is true of $\bar{\mu}_1$. Since (44), where $\mu = \mu^\dagger$, shows that there exists a distribution function $\rho (= \bar{\mu}_2 * \bar{\mu}_3 * \cdots)$ for which Exp $\mu = \bar{\mu}_1 * \rho$, it follows from the case $b = a$ of (25) that Exp μ has an absolutely continuous component. This contradicts the assumption. Hence, $\mu_a \equiv 0$. Consequently, (90) reduces to

$$\text{(90 bis)} \qquad \operatorname{Exp} \mu = (\operatorname{Exp} \mu_d) * (\operatorname{Exp} \mu_s).$$

Hence, it is clear from the case $b = a$ of (25) that neither Exp μ_d nor Exp μ_s has an absolutely continuous component.

Proof of (III). Suppose that Exp μ is purely singular. Then, as just proved, Exp μ_d, Exp μ_s have no absolutely continuous components and satisfy (90 bis). On the other hand, since Exp μ is continuous, it is clear from (90 bis) and (4), that either Exp μ_d or Exp μ_s is continuous. This proves (III).

19. The remarks made at the beginning of § 12 do not preclude for Exp an analogue to a united formulation of the results (I)-(III), § 6 on exp. It turns out that such an analogue actually exists. It lies much deeper than the corresponding fact on exp.

(IV). *If μ is purely discontinuous,* Exp μ *is pure, but can be any of the three types, and it can be absolutely continuous also if $\mu = \mu^\dagger$.*

It is clear from (53) and the case $b = a$ of (24) that, if $\mu \neq \mu^\dagger$, then Exp μ is absolutely continuous. For this reason, (IV) will from now on be restricted to the case $\mu = \mu^\dagger$.

Suppose that $\mu^\dagger$ is purely discontinuous. Then there exist constants p_j, a_j such that (81) holds for every μ'_j in the decomposition (31). Hence, it is seen from (87) that each of the distribution functions (37) is purely discontinuous. But it is known[20] that if $\sigma_1 * \sigma_2 * \cdots$ is any convergent convolution of purely discontinuous distribution functions σ_j, then, as a manifestation of Borel's general principle concerning "either 0 or 1" measures, the distribution function $\sigma_1 * \sigma_2 * \cdots$ is necessarily pure. It follows, therefore, from (44) that Exp $\mu^\dagger$ is pure.

---

[20] B. Jessen and A. Wintner, *loc. cit.*[4], Theorem 35.

In order to complete the proof of (IV), it remains to be shown that Exp $\mu^\dagger$ can be of any of the three types. It is clear from (I), § 12, that Exp $\mu^\dagger$ can be absolutely continuous; for instance, the condition (69) for a smooth Exp $\mu^\dagger$ is satisfied whenever the constants $p_j (\geqq 0)$, $a_j (\neq 0)$ are so chosen that

$$(91_1) \qquad \sum_{j=1}^{\infty} \frac{p_j}{|a_j|^\lambda} = \infty \text{ for some positive } \lambda < 2.$$

On the other hand, (III), § 14, shows that the pure function Exp $\mu^\dagger$ is discontinuous if and only if the numbers $p_j (\geqq 0)$, $a_j (\neq 0)$ are so chosen that

$$(91_2) \qquad \sum_{j=1}^{\infty} \frac{p_j}{a_j^2} < \infty.$$

Finally, the example considered at the end of § 13 was there shown to be continuous but not absolutely continuous and is, therefore, purely singular in view of the theorem just proved.

Incidentally, it will turn out in § 20 bis that

$$(91_3) \qquad S(\text{Exp } \mu^\dagger) = (-\infty, \infty)$$

for the singular Exp μ considered at the end of § 13. Hence, there arises the question whether or not p_j and a_j can so be chosen that the resulting Exp $\mu^\dagger$ is continuous and

$$(91_3 \text{ bis}) \qquad \text{meas } S(\text{Exp } \mu^\dagger) = 0.$$

It will be shown at the end of § 20 bis that the answer is affirmative.

Needless to say, the two sequences $\{a_j\}$, $\{p_j\}$ of numbers determine, by (81) and (31), a distribution function Exp $\mu^\dagger$ if and only if

$$(92) \qquad a_j \gtrless 0, \qquad p_j \geqq 0 \quad \text{and} \quad \sum_{j=1}^{\infty} p_j < \infty,$$

the last series representing the total variation of (31).

20. It is easy to see that, if the p_j are distinct from 0, the spectrum always is

$$(93) \quad S(\text{Exp } \mu^\dagger) = S(\chi_{-c_1}) (+) S_\infty(\chi_{a_1}) (+) (S(\chi_{-c_2}) (+) S_\infty(\chi_{a_2}) (+) \cdots,$$

where the $c = c_j$ are defined by (88).

In fact, it is clear from (9) and from the definition of the set A_∞ in § 1, that, if $A = S(\chi_a)$, the set $S_\infty(\chi_a)$ is the sequence

$$(93 \text{ bis}) \qquad S_\infty(\chi_a): \qquad 0, a, 2a, 3a, \cdots.$$

It is also clear from (9) that $S(p\chi_a)$ is independent of the number $p > 0$; hence, $S_\infty(p\chi_a)$ is identical with the set (93 bis). Since this set is closed, it follows from (81) that (50) reduces to (93).

There is an essential simplification in the symmetric case,

$$a_{2j} = -a_{2j-1} \quad \text{and} \quad p_{2j} = p_{2j-1}; \qquad j = 1, 2, \cdots.$$

In this case, (92) goes over into

$$a_j > 0, p_j > 0 \quad \text{and} \quad \sum_{j=1}^{\infty} p_j < \infty,$$

if the notation is changed in such a way that a_{2j}, p_{2j} become a_j, p_j, respectively. Correspondingly, (88) shows that (93) reduces to

$$S(\operatorname{Exp} \mu) = S_\infty(\chi_{a_1}) (+) S_\infty(\chi_{-a_1}) (+) S_\infty(\chi_{a_2}) (+) S_\infty(\chi_{-a_2}) (+) \cdots.$$

It follows, therefore, from (93 bis) and (48 bis) that

$$\text{(94)} \qquad S(\operatorname{Exp} \mu^\dagger) = \text{closure of all values } \sum_{k=1}^{j} i_k a_k,$$

where i_k denotes an arbitrary real integer and j an arbitrary positive integer. Notice that (94) depends only on the first of the two sequences $\{a_j\}$, $\{p_j\}$.

20 bis. For sake of simplicity, suppose that

$$\sum_{k=1}^{\infty} a_k < \infty; \text{ so that } \sum_{k=1}^{\infty} \pm a_k$$

represents a value for every choice of the signs. It is, therefore, clear from the definitions of the symbols A_∞ and $A\{a_k\}$ in § 1 and § 1 bis, respectively, that

$$A_\infty\{a_k\} = \text{set of all values } \sum_{k=1}^{\infty} i_k a_k, \text{ where } |i_k| \leqq n;$$

it being understood that the integers $n (> 0)$ and $i_k (\gtreqless 0)$ are arbitrary. This obviously implies that

$$\text{closure of } A_\infty\{a_k\} = \text{closure of all values } \sum_{k=1}^{j} i_k a_k, \text{ where } |i_k| \leqq n;$$

it being understood that j and n are arbitrary positive integers, while every $i_k = 0, \pm 1, \pm 2, \cdots$. But the additional restriction, $|i_k| \leqq n$, is now redundant, since it can always be satisfied by choosing, for instance, $n = \max(|i_1|, \cdots, |i_j|)$. It follows, therefore, from (94) that

$$\text{(94 bis)} \qquad S(\operatorname{Exp} \mu^\dagger) = \text{closure of } A_\infty\{a_k\}.$$

It was observed in § 1 bis that if $0 < a < 1$, then $A_\infty\{a^k\}$ is the entire real axis. It follows, therefore, from (94 bis) that (91_3) is satisfied by the continuous function $\operatorname{Exp} \mu^\dagger$ which was constructed at the end of § 13.

On the other hand, it is clear from (94 bis) and from the last remark in § 1 bis, that (91_3 bis) can also be satisfied by a suitable choice of $\{a_k\}$. Then Exp $\mu^\dagger$ has no absolutely continuous component. Hence, Exp $\mu^\dagger$ is purely singular if (91_2) is violated. And (91_2) can be violated by a suitable choice of the $p_j (> 0)$, since the p_j can be chosen arbitrarily within the restriction (92).

APPENDIX.

For the sake of completeness, it will now be explained how the theory of the two exponential operators, as developed in the first three parts of this paper, can be applied to the theory of homogeneous chaoses.

Let there be assigned for every non-negative number t a distribution function σ^t in such a way that

$$\sigma^{t_1} * \sigma^{t_2} = \sigma^{t_1+t_2} \tag{1}$$

and

$$\sigma^t \to \chi_0 \quad \text{as} \quad t \to +0. \tag{1 bis}$$

The condition (1) for the "indefinite divisibility" of $\sigma \equiv \sigma^1$ characterizes the distribution laws by means of which one can introduce into the function spaces of the type considered by Wiener[21] a Jordan content in such a way that the probabilities depend only on the Jordan content, t, of the linear map.[22]

Correspondingly, the continuity axiom (1 bis) concerns the probability carried by sets of small Jordan content and is essential for extension purposes. In fact, while (1) implies only that

$$\sigma^{t_1} * \cdots * \sigma^{t_n} = \sigma^{t_1+\cdots+t_n} \tag{1_n}$$

for every $n < \infty$, it follows from (1 bis) that

$$\sigma^{t_1} * \sigma^{t_2} * \cdots = \sigma^{t_1+t_2+\cdots} \quad \text{whenever} \quad t_1 + t_2 + \cdots < \infty, \tag{1_∞}$$

where the infinite convolution on the left is convergent. One can also say that the rôle of (1 bis) is that of excluding the solutions

$$\sigma^t = \chi_{f(t)} \tag{2}$$

of (1), where $f(t)$ is an arbitrary non-linear (hence, neither measurable nor unilaterally bounded) real-valued solution of the functional equation

$$f(t_1 + t_2) = f(t_1) + f(t_2). \tag{3}$$

The general solution σ^t of (1) was given by Kolmogoroff[23] under the restriction that the distribution functions σ^t possess finite standard deviations. This restriction was subsequently eliminated by Lévy,[24] who found that the Fourier-Stieltjes transform of the most general σ^t is characterized by

---

[21] Cf., e.g., E. Hopf, "Ergodentheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 5, no. 2 (1937), § 16.

[22] Cf. P. Lévy, *loc. cit.*[8]

[23] A. Kolmogoroff, *loc. cit.*[3]

$$\text{(4)} \qquad \log L(u;\sigma^t) = iatu - \tfrac{1}{2}mtu^2 + t\int_{-\infty}^{\infty}\left(e^{iux} - 1 - \frac{iux}{1+x^2}\right) dN(x),$$

where a is an arbitrary real constant, m an arbitrary real non-negative constant, while $N(x)$ is any function which is non-decreasing on both of the intervals $-\infty < x < 0$ and $0 < x < \infty$, remains bounded as $x \to \pm\infty$, and is such that

$$\int_{-1}^{1} x^2 dN(x) < \infty.$$

Actually, (1 bis) is a tacit assumption of Lévy's theorem.[25] In fact, if Zermelo's axiom is admitted in connection with (3), the term $iatu$ in (4) must be replaced by the term $if(t)u$, which corresponds to (2). Incidentally, it can be shown that in this way one obtains the most general solution of (1) alone.

It is easily verified that Lévy's theorem is equivalent to the statement that the general solution σ^t of (1) and (1 bis) is characterized by the existence of an arbitrary real constant c and of an arbitrary non-decreasing bounded function μ for which

$$\sigma^t = \chi_{ct} * \operatorname{Exp}(t\mu^I) * \exp(t\mu^{II}),$$

where the non-decreasing bounded functions μ^I, μ^{II} are defined in terms of μ by placing

$$d\mu^I(x) \equiv \begin{cases} d\mu(x) & \text{if } |x| \leqq 1, \\ 0 & \text{if } |x| > 1, \end{cases} \qquad d\mu^{II}(x) \equiv \begin{cases} 0 & \text{if } |x| \leqq 1, \\ d\mu(x) & \text{if } |x| > 1, \end{cases}$$

and it is understood that

$$\mu^I(-\infty) = 0,\ \mu^I(+\infty) = \mu^I(1); \quad \mu^{II}(-\infty) = 0,\ \mu^{II}(\pm 1 \mp 0) = \mu^{II}(\pm 1)$$

U. S. ARMY,
THE JOHNS HOPKINS UNIVERSITY.

---

[24] P. Lévy, *loc. cit.*[3]

[25] On the other hand, the analogue of (1 bis) in case of infinite convolutions $\sigma_1 * \sigma_2 * \cdots$ is a consequence of the corresponding analogue of (1), that is, of the assumption that the convolution is convergent; cf. B. Jessen and A. Wintner, *loc. cit.*[4] In other words, while the analogue of (1 bis) can be proved, the negation of (1 bis) cannot be disproved. It may be mentioned that this negative formulation of the rôle of Lévy's tacit assumption (1 bis) is not in disagreement with the opinion expressed at the end of his paper "Sur quelques questions de calcul des probabilités," *Prace Matematyczno-Fizyczne*, vol. 39 (1932), pp. 19-28.

ON A PROBLEM OF PARTITIONS.*

By Alfred Brauer.

Introduction. Let $a_1, a_2, \cdots, a_k$ be relatively prime positive integers. The question of determining the number of representations of any integer n in the form

$$(1) \qquad n = a_1x_1 + a_2x_2 + \cdots + a_kx_k \qquad (x_\kappa > 0;\ \kappa = 1, 2, \cdots, k)$$

is one of the problems most treated in the additive theory of numbers. In this paper I shall determine bounds $F(a_1, a_2, \cdots, a_k)$ such that the equation (1) always has solutions in positive integers $x_1, x_2, \cdots, x_k$ for $n > F(a_1, a_2, \cdots, a_k)$. The existence of such a bound is used for instance in the latest researches on the density of the sum of two sets of integers.[1]

If $F(a_1, a_2, \cdots, a_k)$ is such a bound, it is easy to see that

$$G(a_1, a_2, \cdots, a_k) = F(a_1, a_2, \cdots, a_k) - \sum_{\kappa=1}^{k} a_\kappa$$

is a bound for the solubility of (1) in non-negative integers.

For $k = 2$ the solution of the problem has been known for a long time:

$$(2) \qquad F(a_1, a_2) = a_1a_2$$

is such a bound and it is the smallest possible bound; hence it follows that

$$(3) \qquad G(a_1, a_2) = (a_1 - 1)(a_2 - 1) - 1.$$

Sylvester[2] noted that there exist exactly $(a_1 - 1)(a_2 - 1)/2$ integers not representable in the form

$$n = a_1x_1 + a_2x_2 \qquad (x_1 \geqq 0,\ x_2 \geqq 0).$$

For $k > 2$ the problem seems not to be solved; Frobenius mentioned it occasionally in his lectures. In § 1 it will be proved that for $a_1 \leqq a_2 \leqq \cdots \leqq a_k$

$$(4) \qquad G(a_1, a_2, \cdots, a_k) = (a_1 - 1)(a_k - 1) - 1$$

corresponding to (3), and that hence

---

* Received November 25, 1940.

[1] Cf. for instance H. Rohrbach, "Einige neuere Untersuchungen über die Dichte in der additiven Zahlentheorie," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 48 (1938), p. 211.

[2] *Mathematical Questions, with their solutions, from the Educational Times*, vol. 41 (1884), p. 21.

$$(5)\qquad F(a_1, a_2, \cdots, a_k) = a_1 a_k + a_2 + a_3 + \cdots + a_{k-1} = S(a_1, a_2, \cdots, a_k) = S$$

are such bounds. In § 2, the following results will be found instead of (4) and (5): Let

$$d_1 = a_1, \quad d_2 = (a_1, a_2), \cdots, \quad d_k = (a_1, a_2, \cdots, a_k) = 1$$

be the greatest common divisors, then

$$(6)\qquad \begin{aligned} G(a_1, a_2, \cdots, a_k) &= a_2 \frac{d_1}{d_2} + a_3 \frac{d_2}{d_3} + \cdots + a_k \frac{d_{k-1}}{d_k} - \sum_{\kappa=1}^{k} a_\kappa, \\ F(a_1, a_2, \cdots, a_k) &= a_2 \frac{d_1}{d_2} + a_3 \frac{d_2}{d_3} + \cdots + a_k \frac{d_{k-1}}{d_k} = T(a_1, a_2, \cdots, a_k) = T \end{aligned}$$

are such bounds.

In § 3 the bounds (4) and (6) will be compared with each other. It will be proved that $S \geqq T$. We have $S = T$, if and only if S is the least possible bound. This is true if and only if an index r can be found such that $a_2, a_3, \cdots, a_r$ are all divisible by a_1 and such that $a_{r+1} = a_{r+2} = \cdots = a_k$.

In § 4 we inquire as to when T is the best possible bound. It will be proved that this is the case if for $\kappa = 3, 4, \cdots, k$ the integer a_κ / d_κ can be represented in the form

$$(7)\qquad \frac{a_\kappa}{d_\kappa} = \sum_{\nu=1}^{\kappa-1} \frac{a_\nu}{d_{\kappa-1}} y_{\kappa\nu} \quad \text{with} \quad y_{\kappa\nu} \geqq 0.$$

This hypothesis is satisfied, in particular, if

$$a_\kappa > d_{\kappa-1} T(a_1/d_{\kappa-1}, a_2/d_{\kappa-1}, \cdots, a_{\kappa-1}/d_{\kappa-1})$$

holds for $\kappa = 3, 4, \cdots, k$. Therefore for every pair of given integers a_1 and a_2 it is possible to choose $a_3, a_4, \cdots, a_k$ successively such that T is the best bound.

It is interesting that for $k = 3$ the condition (7) is not only sufficient, but also necessary. Set $(a_\kappa, a_\lambda) = d_{\kappa\lambda}$, and

$$a_1 = d_{12} d_{13} b_1, \quad a_2 = d_{12} d_{23} b_2, \quad a_3 = d_{13} d_{23} b_3,$$

and assume $b_1 \leqq b_2 \leqq b_3$. Then we have $(b_1, b_2) = (b_2, b_3) = (b_1, b_3) = 1$, and (7) becomes

$$(8)\qquad b_3 = b_1 y_1 + b_2 y_2 \qquad (y_1 \geqq 0, \ y_2 \geqq 0),$$

and $T(a_1, a_2, a_3)$ is the smallest possible bound if and only if (8) is satisfied. For every pair of relatively prime integers b_1 and b_2 there exist exactly $(b_1 - 1)(b_2 - 1)/2$ integers b_3 for which $T(a_1, a_2, a_3)$ is not the best bound. For instance, $T(m, m+2, m+1)$ is the smallest possible bound if m is an even integer, and not the best bound if $m > 1$ is odd.

For $k > 3$ the condition (7) is not necessary; it can happen that T is the best bound, although (7) is not satisfied. For instance, $T(9, 12, 21, 19)$ is the smallest possible bound.

In § 5 we determine the best bound U for k consecutive integers. It will be proved that

$$U(m, m+1, \cdots, m+k-1) = \left[\frac{m+k^2-3}{k-1}\right] + \frac{k^2-k-2}{2}.$$

The results in § 1 are due to I. Schur; he proved them in his last lecture in Berlin in 1935. The improvement of these results in § 2 is due to the author. The theorems in § 3–5 result partly from discussions of Schur and the author, It was formerly intended to publish these results in a joint paper. I conform with Schur's wishes that the publishing be not longer postponed and that I publish the paper alone.

1. THEOREM 1. *Let* $a_1 \leqq a_2 \leqq \cdots \leqq a_k$ *be relatively prime positive integers. We put*

$$S = S(a_1, a_2, \cdots, a_k) = a_2 + a_3 + \cdots + a_{k-1} + a_1 a_k. \tag{9}$$

For $n > S$ *the Diophantine equation*

$$a_1 x_1 + a_2 x_2 + \cdots + a_k x_k = n \tag{10}$$

always has solutions in positive integers $x_1, x_2, \cdots, x_k$.

Proof. For $k = 2$ the theorem is known since (9) becomes (2). The bound (2) can be obtained easily since when $n = a_1 a_2 + r$ with $r > 0$ the integer x_1 can always be determined such that

$$a_1 x_1 \equiv r \pmod{a_2}$$

where $1 \leqq x_1 \leqq a_2$; hence $x_2 = (n - a_1 x_1)/a_2$ is a positive integer. This result can also be obtained geometrically by using the fact that $(a_1^2 + a_2^2)^{\frac{1}{2}}$ is the distance between two neighboring lattice points on the straight line $a_1 x_1 + a_2 x_2 = n$, where a_1, a_2, n are integers and $(a_1, a_2) = 1$.

Now suppose that $k \geqq 3$. We assume the theorem proved for all the integers less than k. If we put

$$(a_1, a_3, \cdots, a_k) = d, \tag{11}$$

then $(a_2, d) = 1$, and the congruence

$$n \equiv a_2 x_2 \pmod{d} \tag{12}$$

has solutions x_2. We may determine x_2 such that

$$1 \leqq x_2 \leqq d. \tag{13}$$

According to the inductive hypothesis and (12) the equation

$$(14) \qquad \frac{n - a_2x_2}{d} = \frac{a_1}{d}x_1 + \frac{a_3}{d}x_3 + \cdots + \frac{a_k}{d}x_k$$

can be solved in positive integers $x_1, x_3, \cdots, x_k$ if

$$\frac{n - a_2x_2}{d} > \frac{a_3}{d} + \frac{a_4}{d} + \cdots + \frac{a_{k-1}}{d} + \frac{a_1a_k}{d^2}$$

and hence, because of (13), if

$$(15) \qquad n > a_2d + a_3 + a_4 + \cdots + a_{k-1} + \frac{a_1a_k}{d}.$$

Therefore the equations (14) and (10) can be solved in positive integers if (15) is satisfied.

Denote by $S^{(d)}$ the right-hand side of (15). We have only to show that $S^{(d)} \leqq S$. This is true since we have

$$\begin{aligned} S - S^{(d)} &= a_2(1-d) + a_1a_k\frac{d-1}{d} \\ &= (d-1)\left(\frac{a_1a_k}{d} - a_2\right) \geqq (d-1)(a_k - a_2) \geqq 0 \end{aligned}$$

by (9), (15), and (11).

COROLLARY. *For* $n \geqq (a_1 - 1)(a_k - 1)$ *the equation* (10) *always has solutions in non-negative integers.*

2. Without using Theorem 1 we shall now obtain another bound.

THEOREM 2. *Let* $a_1, a_2, \cdots, a_k$ *be relatively prime positive integers. If we put*

$$(a_1, a_2, \cdots, a_\kappa) = d_\kappa \qquad (\kappa = 1, 2, \cdots, k)$$

the Diophantine equation (10) *always has solutions in positive integers* $x_1, x_2, \cdots, x_k$ *if*

$$n > a_2\frac{d_1}{d_2} + a_3\frac{d_2}{d_3} + \cdots + a_k\frac{d_{k-1}}{d_k} = T(a_1, a_2, \cdots, a_k) = T$$

holds.

By changing the numbering of the a_κ the value of T is possibly changed.

Proof. Since a_1/d_2 and a_2/d_2 are relatively prime, every integer m greater than a_1a_2/d_2^2 can be represented in the form

$$m = \frac{a_1}{d_2}y_1 + \frac{a_2}{d_2}y_2 \qquad (y_1 > 0,\ y_2 > 0).$$

Therefore every integer n divisible by d_2 and greater than

$$a_1a_2/d_2 = a_2d_1/d_2$$

is representable in the form

$$n = a_1x_1 + a_2x_2 \qquad (x_1 > 0,\ x_2 > 0).$$

Now we shall prove for $\kappa = 2, 3, \cdots, k$ that every integer N satisfying both the conditions

(16) $$N \equiv 0 \pmod{d_\kappa}$$

and

(17) $$N > a_2\frac{d_1}{d_2} + a_3\frac{d_2}{d_3} + \cdots + a_\kappa\frac{d_{\kappa-1}}{d_\kappa}$$

can be represented in the form

(18) $$N = a_1x_1 + a_2x_2 + \cdots + a_\kappa x_\kappa \qquad (x_1 > 0, x_2 > 0, \cdots, x_\kappa > 0).$$

For $\kappa = 2$ this result has been already proved. We assume it to be true for $\kappa = 2, 3, \cdots, l-1$, and we shall prove that it is true for $\kappa = l$. Because of (17) and (16) we can set

(19) $$N = a_2\frac{d_1}{d_2} + a_3\frac{d_2}{d_3} + \cdots + a_l\frac{d_{l-1}}{d_l} + bd_l \quad \text{with} \quad b \geqq 1.$$

The congruence

(20) $$a_lx_l \equiv bd_l \pmod{d_{l-1}}$$

has a solution x_l since

(21) $$(a_l, d_{l-1}) = d_l.$$

It follows from (20) and (21) that

$$x_l \equiv \frac{bd_l}{a_l}\left(\operatorname{mod} \frac{d_{l-1}}{d_l}\right).$$

Hence we can assume that

(22) $$0 < x_l \leqq \frac{d_{l-1}}{d_l}.$$

Therefore we obtain from (22) and (19)

(23) $$N - a_lx_l \geqq N - a_l\frac{d_{l-1}}{d_l} > a_2\frac{d_1}{d_2} + a_3\frac{d_2}{d_3} + \cdots + a_{l-1}\frac{d_{l-2}}{d_{l-1}}$$

and from (19) and (20)

(24) $$N - a_lx_l = a_2\frac{d_1}{d_2} + a_3\frac{d_2}{d_3} + \cdots + a_{l-1}\frac{d_{l-2}}{d_{l-1}} + \frac{a_l}{d_l}d_{l-1} + bd_l - a_lx_l \equiv 0 \pmod{d_{l-1}}.$$

According to the inductive hypothesis it follows from (23) and (24) that $N - a_lx_l$ can be represented in the form

$$N - a_l x_l = a_1 x_1 + a_2 x_2 + \cdots + a_{l-1} x_{l-1} \qquad (x_\lambda > 0 \text{ for } \lambda = 1, 2, \cdots, l-1)$$

whence

$$N = a_1 x_1 + a_2 x_2 + \cdots + a_{l-1} x_{l-1} + a_l x_l \qquad (x_\lambda > 0 \text{ for } \lambda = 1, 2, \cdots, l).$$

Thus, if (16) and (17) are satisfied, (18) is true for every κ, hence also for $\kappa = k$. This proves Theorem 2.

3. We now study the question as to when the bounds S and T are the smallest possible.

THEOREM 3. *The bound S is the smallest possible if an index r can be found such that*

$$a_2 \equiv a_3 \equiv \cdots \equiv a_r \equiv 0 \pmod{a_1} \tag{25}$$

and

$$a_{r+1} = a_{r+2} = \cdots = a_k. \tag{26}$$

Proof. Writing $k - r = m$ we get

$$S = a_2 + a_3 + \cdots + a_r + (m-1)a_k + a_1 a_k \tag{27}$$

because of (26). If S is representable in the form $\sum_{\nu=1}^{k} a_\nu x_\nu$ with $x_\nu \geqq 1$, we have

$$S = a_1 x_1 + a_2 x_2 + \cdots + a_r x_r + a_k z \tag{28}$$

where

$$z \geqq m. \tag{29}$$

By (27) and (25) we get

$$S \equiv (m-1)a_k \pmod{a_1}$$

and by (28)

$$S \equiv a_k z \pmod{a_1}$$

hence

$$a_k z \equiv (m-1)a_k \pmod{a_1},$$
$$z \equiv m - 1 \pmod{a_1}$$

since a_1 and a_k are relatively prime. Therefore we have

$$z \geqq m - 1 + a_1$$

because of (29). It follows from (28) and (27) that

$$S = a_1 x_1 + a_2 x_2 + \cdots + a_r x_r + a_k z \geqq a_1 + a_2 + \cdots + a_r + (m-1)a_k + a_1 a_k = S + a_1.$$

This is not possible; hence S is not representable in the form $\sum_{\nu=1}^{k} a_\nu x_\nu$ with $x_\nu \geqq 1$, and the theorem is proved.

If the conditions (25) and (26) are not satisfied, however, S is not the best bound:

THEOREM 4. *We have* $S \geqq T$. *Here the sign of equality is true if and only if the conditions* (25) *and* (26) *are satisfied.*

Proof. We have

$$S - T = a_2 + a_3 + \cdots + a_{k-1} + a_1 a_k$$
$$- a_2 \frac{d_1}{d_2} - a_3 \frac{d_2}{d_3} - \cdots - a_{k-1} \frac{d_{k-2}}{d_{k-1}} - a_k d_{k-1},$$

nce

$$0) \quad S - T = a_k(a_1 - d_{k-1}) - \frac{a_2}{d_2}(d_1 - d_2) - \frac{a_3}{d_3}(d_2 - d_3) - \cdots - \frac{a_{k-1}}{d_{k-1}}(d_{k-2} - d_{k-1})$$
$$\geqq a_k(a_1 - d_{k-1} - d_1 + d_2 - d_2 + d_3 - + \cdots - d_{k-2} + d_{k-1}) = 0.$$

It is now a question as to when $S = T$ holds. We have to distinguish between two cases.

1) In the first place we assume that

$$d_1 = d_2 = \cdots = d_{k-1}. \tag{31}$$

Then $a_1 = d_1 = d_{k-1}$; all the parentheses in (30) are zero, and it follows that $S = T$. Since $a_1 = d_1$ and $d_{k-1} = (a_1, a_2, \cdots, a_{k-1})$ condition (31) is satisfied if

$$a_2 \equiv a_3 \equiv \cdots \equiv a_{k-1} \equiv 0 \pmod{a_1}.$$

Therefore (25) and (26) are satisfied for $r = k - 1$.

2) Next we assume that

$$d_1 \neq d_{k-1}.$$

Let κ be the smallest index such that $d_{\kappa-1} > d_\kappa$. From $d_{\kappa-1} - d_\kappa > 0$ it follows that the sign of equality in (30) is possible only if

$$a_\kappa / d_\kappa = a_k. \tag{32}$$

Since $a_\kappa \leqq a_k$, we get $d_\kappa = 1$ from (32) and we have $a_\kappa = a_{\kappa+1} = \cdots = a_k$. Hence (26) is satisfied for $r = \kappa - 1$. Since κ was the smallest index of this kind, it follows that

$$a_1 = d_1 = d_2 = \cdots = d_{\kappa-1}$$

and therefore that $d_{\kappa-1}$ is divisible by a_1, and that

$$a_2 \equiv a_3 \equiv \cdots \equiv a_{\kappa-1} \equiv 0 \pmod{a_1}.$$

Thus (25) is also satisfied for $r = \kappa - 1$.

Corollary. *The bound S is the best possible if and only if $S = T$.*

4. Now we shall inquire as to when the bound T is the best possible.

Theorem 5. *Suppose that $k \geqq 3$. If for $\kappa = 3, 4, \cdots, k$ the integer a_κ/d_κ is representable in the form*

$$\frac{a_\kappa}{d_\kappa} = \sum_{\nu=1}^{\kappa-1} \frac{a_\nu}{d_{\kappa-1}} y_{\kappa\nu} \quad \text{with} \quad y_{\kappa\nu} \geqq 0, \tag{33}$$

then $T(a_1, a_2, \cdots, a_k)$ is the smallest possible bound.

The hypotheses of this theorem are satisfied in particular if

$$a_\kappa/d_\kappa > T(a_1/d_{\kappa-1}, a_2/d_{\kappa-1}, \cdots, a_{\kappa-1}/d_{\kappa-1}) \quad \text{for} \quad \kappa = 3, 4, \cdots, k,$$

for then a_κ/d_κ is certainly representable in the form (33), and even with positive $y_{\kappa\nu}$. Since $d_\kappa \leqq d_{\kappa-1}$, it is easy to determine integers $a_3, a_4, \cdots, a_k$ when a_1 and a_2 are given such that for these integers $T(a_1, a_2, \cdots, a_k)$ is the smallest possible bound.

Proof. We shall prove that $T(a_1/d_\kappa, a_2/d_\kappa, \cdots, a_\kappa/d_\kappa)$ is the best bound for $a_1/d_\kappa, a_2/d_\kappa, \cdots, a_\kappa/d_\kappa$ if (33) holds. For $\kappa = 2$ the bound T is always the best possible. Suppose that the statement is already proved for $2, 3, \cdots, \kappa - 1$. Then the integer $T(a_1/d_{\kappa-1}, a_2/d_{\kappa-1}, \cdots, a_{\kappa-1}/d_{\kappa-1})$ is not representable in the form $\sum_{\nu=1}^{\kappa-1} (a_\nu/d_{\kappa-1}) x_\nu$ with $x_\nu > 0$ and we have to prove that $T(a_1/d_\kappa, a_2/d_\kappa, \cdots, a_\kappa/d_\kappa)$ is also not representable in the form $\sum_{\nu=1}^{k} (a_\nu/d_\kappa) x_\nu$ with $x_\nu > 0$. If this were not true, we would have

$$T\left(\frac{a_1}{d_\kappa}, \frac{a_2}{d_\kappa}, \cdots, \frac{a_\kappa}{d_\kappa}\right) = \frac{a_2 a_1}{d_\kappa d_2} + \frac{a_3 d_2}{d_\kappa d_3} + \cdots + \frac{a_\kappa d_{\kappa-1}}{d_\kappa^2} = \frac{a_1}{d_\kappa} x_1 + \frac{a_2}{d_\kappa} x_2 + \cdots + \frac{a_\kappa}{d_\kappa} x_\kappa \tag{34}$$

where $x_1 > 0, x_2 > 0, \cdots, x_\kappa > 0$, and thus

$$\frac{a_1 a_2}{d_2} + \frac{a_3 d_2}{d_3} + \cdots + \frac{a_\kappa d_{\kappa-1}}{d_\kappa} = a_1 x_1 + a_2 x_2 + \cdots + a_\kappa x_\kappa. \tag{35}$$

Therefore, since d_κ is the greatest common divisor of a_κ and $d_{\kappa-1}$, we would have

$$a_\kappa x_\kappa \equiv 0 \pmod{d_{\kappa-1}},$$
$$x_\kappa \equiv 0 \pmod{d_{\kappa-1}/d_\kappa}.$$

Putting $x_\kappa = x'_\kappa d_{\kappa-1}/d_\kappa$ we get from (35) and (33)

$$(36)\qquad \frac{a_1a_2}{d_2}+\frac{a_3d_2}{d_3}+\cdots+\frac{a_{\kappa-1}d_{\kappa-2}}{d_{\kappa-1}}$$
$$= a_1x_1+a_2x_2+\cdots+a_{\kappa-1}x_{\kappa-1}+a_\kappa x'_\kappa\frac{d_{\kappa-1}}{d_\kappa}-a_\kappa\frac{d_{\kappa-1}}{d_\kappa}$$
$$= a_1x_1+a_2x_2+\cdots+a_{\kappa-1}x_{\kappa-1}+(x'_\kappa-1)\sum_{\nu=1}^{\kappa-1}a_\nu y_{\kappa\nu}$$
$$= a_1z_1+a_2z_2+\cdots+a_{\kappa-1}z_{\kappa-1}$$

where $z_1, z_2, \cdots, z_{\kappa-1}$ are positive integers. On the other hand we have

$$(37)\quad T\left(\frac{a_1}{d_{\kappa-1}},\frac{a_2}{d_{\kappa-1}},\cdots,\frac{a_{\kappa-1}}{d_{\kappa-1}}\right)=\frac{a_1a_2}{d_{\kappa-1}d_2}+\frac{a_3d_2}{d_{\kappa-1}d_3}+\cdots+\frac{a_{\kappa-1}d_{\kappa-2}}{d^2_{\kappa-1}}.$$

From (36) and (37) we get

$$T\left(\frac{a_1}{d_{\kappa-1}},\frac{a_2}{d_{\kappa-1}},\cdots,\frac{a_{\kappa-1}}{d_{\kappa-1}}\right)=\frac{a_1}{d_{\kappa-1}}z_1+\frac{a_2}{d_{\kappa-1}}z_2+\cdots+\frac{a_{\kappa-1}}{d_{\kappa-1}}z_{\kappa-1}.$$

This is incompatible with the inductive hypothesis, and therefore the hypothesis (34) is not possible. Thus the assumption is proved for every κ, and for $\kappa = k$ we get Theorem 5.

Theorem 5 gives a sufficient condition that T be the smallest possible bound. In the case $k = 3$ we shall show that this condition is also necessary. For this purpose we shall prove two lemmas.

LEMMA 1. *Let a and b be relatively prime positive integers. Then every positive integer m not divisible by a or by b is representable either in the form*

$$(38)\qquad m = ax + by \qquad (x > 0,\ y > 0)$$

or in the form

$$(39)\qquad m = ab - ax - by \qquad (x > 0,\ y > 0).$$

Proof. Let u be a solution of the congruence

$$au \equiv m \pmod b$$

such that $0 < u < b$. If we put

$$m = au + bv$$

the integer m is represented in the form (38) if v is positive. If v is negative, however, $v = -y$ with $y > 0$, we obtain

$$m = ab - (ab - au) - by = ab - ax - by$$

where $x = b - u > 0$; thus m is represented in the form (39).

Finally m is not representable simultaneously in the forms (38) and (39), because if it were, we would have $m < ab$ and

$$ax + by = ab - ax' - by' \tag{40}$$

It would then follow from (40) that

$$ab = a(x + x') + b(y + y').$$

This is impossible.

COROLLARY 1. *Let a and b be relatively prime positive integers. There are $(a-1)(b-1)/2$ integers of the set $1, 2, \cdots, ab$ representable in the form* (38) *and just as many in the form* (39).

Proof. The multiples of a and the multiples of b less than or equal to ab are not representable by (38) or by (39). Let m be one of the remaining $(a-1)(b-1)$ positive integers less than ab. If m is representable in the form (38), then $ab - m$ is representable in the form (39).

COROLLARY 2 (SYLVESTER). *If we have $(a, b) = 1$, then $(a-1)(b-1)/2$ integers are not representable in the form*

$$m = ax + by \quad \textit{with} \quad x \geqq 0,\ y \geqq 0. \tag{41}$$

Proof. The multiples of a and those of b are representable in the form (41).

LEMMA 2. *Let a, b, m be positive integers, $(a, b) = 1$, and m divisible neither by a nor by b. If m is representable in the form*

$$m = ab - ax - by \qquad (x > 0, y > 0), \tag{42}$$

then every divisor of m is representable in the same form.

Proof. Let t be a divisor of m and put $m = tm_1$. If t is not representable in the form (42) then t is representable in the form

$$t = au + bv$$

by Lemma 1 since t is neither divisible by a nor by b. Hence

$$m = tm_1 = am_1u + bm_1v.$$

This is impossible because of (42), and therefore t is representable in the form (42).

Now let a_1, a_2, a_3 be three relatively prime positive integers. We put

$$\begin{cases} (a_\kappa, a_\lambda) = d_{\kappa\lambda} & (\kappa, \lambda = 1, 2, 3;\ \kappa \neq \lambda), \\ a_1 = d_{12}d_{13}b_1, \quad a_2 = d_{12}d_{23}b_2, & a_3 = d_{13}d_{23}b_3 \end{cases} \tag{43}$$

where

$$(b_1, b_2) = (b_1, b_3) = (b_2, b_3) = 1. \tag{44}$$

Let us assume for instance that $b_1 \leqq b_2 \leqq b_3$. Then we have

$$b_1b_2 + b_3 \leqq \min(b_1b_3 + b_2,\ b_2b_3 + b_1) \tag{45}$$

and

$$(46)\qquad T = T(a_1, a_2, a_3) = a_2 a_1/d_{12} + a_3 d_{12} = d_{12}d_{13}d_{23}(b_1 b_2 + b_3) \leqq \min\{T(a_1, a_3, a_2), T(a_2, a_3, a_1)\}$$

because of (45). Here the condition (33) has the form

$$a_3 = (a_1/d_{12})y_{31} + (a_2/d_{12})y_{32} \qquad (y_{31} \geqq 0,\ y_{32} \geqq 0),$$

hence

$$d_{13}d_{23}b_3 = d_{13}b_1 y_{31} + d_{23}b_2 y_{32}.$$

By (43) and (44) we have

$$(d_{13}, d_{23}b_2) = (d_{23}, d_{13}b_1) = 1,$$

and therefore y_{32} is divisible by d_{13} and y_{31} by d_{23}. Dividing by $d_{13}d_{23}$ we observe that (33) has the form

$$(47)\qquad b_3 = b_1 y_1 + b_2 y_2 \qquad (y_1 \geqq 0,\ y_2 \geqq 0).$$

We shall now prove that (47) is also necessary in the case $k = 3$. Moreover, we shall give a second proof for Theorem 5 in the case $k = 3$.

THEOREM 6. *For $k = 3$ the bound T is the best possible if and only if (47) holds.*

Proof. If T is representable in the form

$$(48)\qquad T = a_1 x_1 + a_2 x_2 + a_3 x_3 \qquad (x_1 > 0,\ x_2 > 0,\ x_3 > 0),$$

then T is not the smallest bound. Therefore we have to prove that T is representable in the form (48) if and only if (47) is not satisfied.

Suppose that (47) is not satisfied. This is only possible if $b_1 > 1$. It then follows from Lemma 1 that

$$b_3 = b_1 b_2 - y'_1 b_1 - y'_2 b_2 \quad \text{with} \quad y'_1 > 0,\ y'_2 > 0,$$

and hence by (46) and (43)

$$T = d_{12}d_{13}d_{23}(b_1 b_2 + b_3) = d_{12}d_{13}d_{23}(b_3 + y'_1 b_1 + y'_2 b_2 + b_3) = d_{23}y'_1 a_1 + d_{13}y'_2 a_2 + 2d_{12}a_3;$$

therefore T is representable in the form (48).

Conversely, suppose that (48) is satisfied. Then we obtain from (46)

$$d_{12}d_{13}d_{23}b_1 b_2 + d_{12}d_{13}d_{23}b_3 = T = d_{12}d_{13}b_1 x_1 + d_{12}d_{23}b_2 x_2 + d_{13}d_{23}b_3 x_3.$$

As previously, it follows that x_1 is divisible by d_{23} and, in the same way, that x_2 is divisible by d_{13}, and x_3 by d_{12}. Therefore we have

$$(49)\qquad b_1 b_2 + b_3 = b_1 x'_1 + b_2 x'_2 + b_3 x'_3$$

where x'_1, x'_2, x'_3 are positive integers. This cannot happen if $b_1 = 1$. For $b_1 > 1$ we have $x'_3 > 1$ since

$$b_1 b_2 = b_1 x'_1 + b_2 x'_2$$

is impossible because of $(b_1, b_2) = 1$. Hence we obtain from (49)

$$b_3(x'_3 - 1) = b_1 b_2 - b_1 x'_1 - b_2 x'_2$$

where $x'_3 - 1 \geqq 1$. It now follows from Lemma 2 that b_3 is representable in the form

$$b_3 = b_1 b_2 - b_1 x''_1 - b_2 x''_2 \qquad (x''_1 > 0,\ x''_2 > 0),$$

and from Lemma 1 that (47) is not satisfied.

COROLLARY 1. *Let* b_1 *and* b_2 *be given integers. Then there exist exactly* $(b_1 - 1)(b_2 - 1)/2$ *positive integers* b_3 *for which* $T(a_1, a_2, a_3)$ *is not the best bound.*

This follows from the Corollary 2 of Lemma 1.

COROLLARY 2. *Let* $a_1 = m$, $a_2 = m + 1$, $a_3 = m + 2$ *be three consecutive integers with* $m > 1$. *The bound* $T(m, m + 2, m + 1)$ *is the best possible bound if and only if* m *is even.*

Proof. If m is even, we have

$$d_{13} = 2, \quad d_{12} = d_{23} = 1; \quad b_1 = \tfrac{1}{2}m, \quad b_2 = m + 1, \quad b_3 = \tfrac{1}{2}m + 1,$$

and hence $b_2 = b_1 + b_3$. Thus the condition (47) is satisfied and T is the best bound. If m is odd, $m, m + 1, m + 2$ are relatively prime in pairs, and (47) is not satisfied. In the latter case, we shall obtain the smallest possible bound in **5**.

5. THEOREM 7. *Denote by* $U(m, m + 1, \cdots, m + k - 1)$ *the smallest possible bound for* k *consecutive integers* $m, m + 1, \cdots, m + k - 1$. *Then we have*

$$U(m, m + 1, \cdots, m + k - 1) = \left[\frac{m + k^2 - 3}{k - 1}\right] m + \frac{k^2 - k - 2}{2}.$$

Proof. Let N be any positive integer representable in the form

$$N = m x_1 + (m + 1) x_2 + \cdots + (m + k - 1) x_k \qquad (x_\kappa > 0;\ \kappa = 1, 2, \cdots, k). \tag{50}$$

We set

$$\left\{\begin{array}{l} x_1 + x_2 + \cdots + x_k = y, \\ x_2 + 2x_3 + \cdots + (k - 1)x_k = z, \end{array}\right. \tag{51}$$

and accordingly

$$N = my + z. \tag{52}$$

Here we have $y \geqq k$ because of $x_\kappa > 0$; moreover it follows from (51) that

$$\binom{k}{2} \leqq z = (k-1)y - x_{k-1} - 2x_{k-2} - \cdots - (k-1)x_1$$
$$\leqq (k-1)y - \binom{k}{2} = \binom{k}{2} + (k-1)(y-k).$$

It is easy to see that for every fixed y the integer z can actually take each of the values

$$\binom{k}{2},\ \binom{k}{2}+1, \cdots, \binom{k}{2} + (k-1)(y-k).$$

For, if z is one of these integers, we divide the non-negative integer $z - \frac{1}{2}k(k-1)$ by $k-1$:

$$z - \binom{k}{2} = (k-1)q + r \quad \text{with} \quad 0 \leqq r < k-1.$$

Then it is sufficient to choose

$$x_\kappa = 1 \quad \text{for} \quad 2 \leqq \kappa \leqq k-1, \quad \kappa \neq r+1$$

and

$$x_k = q+1, \quad x_{r+1} = 2, \quad x_1 = y-k-q \quad \text{for} \quad r > 0,$$
$$x_k = q+1, \quad x_1 = y-k-q+1 \quad \text{for} \quad r = 0.$$

This is possible since

$$(k-1)q + r = z - \binom{k}{2} \leqq (k-1)(y-k),$$

and hence

$$y-k-q \geqq 1 \quad \text{for} \quad r > 0, \quad \text{and} \quad y-k-q+1 \geqq 1 \quad \text{for} \quad r = 0.$$

In both cases the conditions (51) are satisfied.

If we set $g = km + \frac{1}{2}k(k-1)$, then it follows from (52) that the integers

$$(53)\quad \begin{cases} g \\ g+m,\ g+m+1, \ \ldots, g+m+k-1 \\ g+2m,\ g+2m+1, \ \ldots, g+2m+2(k-1) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ g+(y-k)m, \ \ldots, g+(y-k)m+(y-k)(k-1) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{cases}$$

form the totality of positive integers representable in the form (50). The first elements of two successive lines in (53) have the difference m. The line beginning with $g+(y-k)m$ contains exactly $(y-k)(k-1)+1$ consecutive integers. Therefore the first line from which on all the following integers are representable must contain not less than m integers. It is necessary for this that y is the smallest integer for which

$$m \leqq (y-k)(k-1)+1; \quad y \geqq \frac{m-1}{k-1}+k. \tag{54}$$

Since y is the smallest integer of this kind, we have

$$y = \left[\frac{m-1+k-2}{k-1}\right]+k = \left[\frac{m+k^2-3}{k-1}\right]. \tag{55}$$

The preceding line contains only $(y-1-k)(k-1)+1$ integers, hence it contains less than m integers because of (54). Therefore it follows from (55) that the greatest integer not representable is

$$N = g+(y-k)m-1 = \left[\frac{m+k^2-3}{k-1}\right]m+\frac{k^2-k-2}{2}.$$

This proves Theorem 7.

In particular, we get for $k=3$

$$\begin{aligned} U(m, m+1, m+2) &= \left[\frac{m+6}{2}\right]m+2 = \frac{m^2+6m+4}{2} \\ &= T(m, m+2, m+1) \end{aligned}$$

if m is even, and

$$\begin{aligned} U(m, m+1, m+2) &= \left[\frac{m+6}{2}\right]m+2 = \frac{(m+5)m}{2}+2 \\ &= \frac{m^2+5m+4}{2} < m^2+2m+2 = T(m, m+1, m+2) \end{aligned}$$

if $m>1$ is odd.

Institute for Advanced Study,
Princeton, N. J.

ON TWO CONJECTURES IN THE THEORY OF NUMBERS.*

By A. E. INGHAM.

---

1. Let

$$M(x) = \sum_{n \leq x} \mu(n), \qquad L(x) = \sum_{n \leq x} \lambda(n),$$

where $\mu(n)$ and $\lambda(n)$ are the arithmetical functions of Möbius and Liouville defined by

$$(1) \qquad 1/\zeta(s) = \prod_p (1 - (1/p^s)) = \sum_{n=1}^{\infty} \mu(n)/n^s$$

$$(2) \qquad \zeta(2s)/\zeta(s) = \prod_p [1 - (1/p^s) + (1/p^{2s}) - \cdots] = \sum_{n=1}^{\infty} \lambda(n)/n^s$$

$$(s = \sigma + it, \sigma > 1).$$

It was suggested by Stieltjes [6] and by Mertens [1] that $|M(x)| < x^{1/2}$ $(x > 1)$. The calculations of von Sterneck [3, 4, 5] confirm this up to $x = 500{,}000$, and at intervals up to 5,000,000, and indeed with a factor $1/2$ on the right hand side except near $x = 200$.

It was conjectured by Pólya [2] that $L(x) \leq 0$ $(x \geq 2)$. This is supported by a smaller body of numerical evidence (Pólya verified it up to $x = 1{,}500$),[1] but the conjecture is attractive on account of its connection with the theory of binary quadratic forms $ax^2 + bxy + cy^2$ (a, b, c integers). Pólya showed that a positive integer m is a solution of the equation $L(m) = 0$ whenever $4m + 3$ is a prime $p > 7$ for which $h(-p) = 1$, where $h(D)$ is the number of classes of (positive definite) forms of (negative) discriminant $D(= b^2 - 4ac)$. If these were the only solutions of $L(m) = 0$, it would be easy to deduce, in view of the fact that, since $\lambda(n) = \pm 1$, $L(x)$ cannot change sign without vanishing, and of Heilbronn's theorem that $h(D) = 1$ for only a finite number of negative D, that $L(x) < 0$ at any rate for all sufficiently large x. But other solutions of $L(m) = 0$ occur within the limits of Pólya's calculations, and the argument decides nothing.

It is well known (and the proof is reproduced for completeness at the beginning of § 3) that the truth of either of the above conjectures, or more generally the truth of any one of the four inequalities

$$M(x) < Kx^{1/2}, \quad M(x) > -Kx^{1/2}, \quad L(x) < Kx^{1/2}, \quad L(x) > -Kx^{1/2},$$

---

* Received February 12, 1941.

[1] The verification has recently been extended to $x = 20{,}000$. See H. Gupta, "On a table of values of $L(n)$," *Proc. Indian Acad. Sci.*, Sect. A. vol 12 (1940), pp. 407-409.

for all sufficiently large x, where K is a constant, would imply that $\zeta(s)$ has all its complex zeros on the line $\sigma = \frac{1}{2}$ (i. e., that the Riemann hypothesis is true), and that all the zeros are simple. The purpose of this note is to point out a further consequence, namely that the imaginary parts of the zeros above the real axis must be linearly dependent (with rational integral multipliers). More precisely, we shall prove

THEOREM A. *If the imaginary parts* $\gamma_1, \gamma_2, \cdots$ *of the (distinct) zeros of* $\zeta(s)$ *above the real axis are connected by no relation of the type*

$$\sum_{n=1}^{N} c_n\gamma_n = 0 \quad (c_n \textit{ integers not all } 0), \tag{3}$$

or by only a finite number of such relations, then, when $x \to \infty$,

$$\overline{\lim}\, x^{-\frac{1}{2}}M(x) = +\infty, \qquad \underline{\lim}\, x^{-\frac{1}{2}}M(x) = -\infty,$$
$$\overline{\lim}\, x^{-\frac{1}{2}}L(x) = +\infty, \qquad \underline{\lim}\, x^{-\frac{1}{2}}L(x) = -\infty.$$

The method of proof is similar to that by which Littlewood disproved the conjecture $\pi(x) < \operatorname{li} x$ in the theory of primes, except that the use of Phragmén-Lindelöf theorems and of 'explicit formulae' is avoided by an application of the technique introduced by Wiener in his fundamental researches on Tauberian theorems, and that Dirichlet's theorem on Diophantine approximation is replaced by Kronecker's theorem. It is this that calls for the hypothesis of linear independence of the γ_n. It would be easy to relax this hypothesis a little, but there seems no obvious way of replacing it by anything essentially easier to verify.

2. The proof of Theorem A is based on a theorem concerning Laplace integrals (Theorem 1) and on a special property of the generating functions (1) and (2) (Theorem 2).

THEOREM 1. *Let*

$$F(s) = \int_0^\infty A(u)e^{-su}du, \tag{4}$$

where $A(u)$ *is absolutely integrable over every finite interval* $0 \leq u \leq U$, *and the integral is convergent in some half-plane* $\sigma > \sigma_1 \geq 0$.

Let $A^*(u)$ *be a real trigonometrical polynomial*

$$A^*(u) = \sum_{-N}^{N} \alpha_n e^{i\gamma_n u} \qquad (\gamma_n \textit{ real}, \gamma_{-n} = -\gamma_n, \alpha_{-n} = \bar{\alpha}_n),$$

and let

$$F^*(s) = \int_0^\infty A^*(u)e^{-su}du = \sum_{-N}^{N} \frac{\alpha_n}{s - i\gamma_n} \qquad (\sigma > 0).$$

Suppose that $F(s) - F^*(s)$ (*suitably defined outside the half-plane* $\sigma > \sigma_1$) *is regular in the region* $\sigma \geq 0$, $-T \leq t \leq T$, *for some* $T > 0$ (*or, more generally, continuous in this region and regular in the interior*).

Then, when $u \to \infty$ (*T fixed*)

$$\underline{\lim}\, A(u) \leq \underline{\lim}\, A^*_T(u) \leq \overline{\lim}\, A(u), \tag{5}$$

where

$$A^*_T(u) = \sum_{|\gamma_n| < T} [1 - (|\gamma_n|/T)]\alpha_n e^{i\gamma_n u} = \alpha_0 + 2\Re \sum_{0 < \gamma_n < T} [1 - (\gamma_n/T)]\alpha_n e^{i\gamma_n u}.$$

It is enough to prove the inequality for the lower limits, since the other may be deduced by changing signs. We may suppose that $\underline{\lim}\, A(u) > -\infty$ (since otherwise there is nothing to prove), and indeed that $\underline{\lim}\, A(u) > 0$ [by adding a constant to $A(u)$ and to $A^*(u)$]. Then $A(u) > 0$ for $u > u_0$ (say), and it follows from a well-known theorem of Landau that the integral in (4) is convergent, and the relation (4) true, for $\sigma > 0$, since $F(s)$ is regular along the positive real axis $s > 0$, this being true of $F^*(s)$ (obviously) and of $D(s) = F(s) - F^*(s)$ (by hypothesis). Thus we have (on our suppositions)

$$D(s) = \int_0^\infty A(u)e^{-su}du - \int_0^\infty A^*(u)e^{-su}du \qquad (\sigma > 0). \tag{6}$$

Let $\quad k(t) = k_T(t) = 1 - (|t|/T) \quad (|t| < T); \quad = 0 \quad (|t| \geq T)$,

$$K(v) = K_T(v) = T[(\sin Tv/2)/(Tv/2)]^2,$$

so that

$$\int_{-T}^{T} k(t)e^{-itv}dt = K(v), \qquad \int_{-\infty}^{\infty} K(v)e^{itv}dv = 2\pi k(t).$$

Multiplying (6) by $k(t)e^{it\omega}$ $(\omega > 0)$ and integrating over $-T \leq t \leq T$, we obtain [since the integrals in (6) are uniformly convergent over this range for fixed $\sigma > 0$]

$$\int_{-T}^{T} D(\sigma + it)k(t)e^{it\omega}dt = \int_0^\infty A(u)K(u - \omega)e^{-\sigma u}du$$
$$- \int_0^\infty A^*(u)K(u - \omega)e^{-\sigma u}du \qquad (\sigma > 0).$$

The first and third integrals here are continuous functions of σ for $\sigma \geq 0$, the first because $D(s)$ is continuous for $\sigma \geq 0$, $-T \leq t \leq T$, and the third because the integral is (absolutely) convergent when $\sigma = 0$ [the integrand then being $O(u^{-2})$ as $u \to \infty$ (T and ω fixed)]. The second integral, having a non-negative integrand for $u > u_0$, is continuous for $\sigma \geq 0$ or tends to $+\infty$

when $\sigma \to +0$ (according as it is convergent or divergent to $+\infty$ when $\sigma = 0$). The latter alternative is, however, excluded by the behaviour of the other two integrals, and we obtain, on making $\sigma \to +0$,

$$\int_{-T}^{T} D(it)k(t)e^{it\omega}dt = \int_{0}^{\infty} A(u)K(u-\omega)\,du - \int_{0}^{\infty} A^*(u)K(u-\omega)\,du.$$

Now let $\omega \to \infty$. The first integral tends to 0 by the Riemann-Lebesgue theorem, and we obtain

$$(7) \qquad \int_{0}^{\infty} A(u)K(u-\omega)\,du = \int_{0}^{\infty} A^*(u)K(u-\omega)\,du + o(1)$$

as $\omega \to \infty$ (T fixed).

Since $K(v) \geq 0$ the first integral here is

$$\geq -\max_{-\omega \leq v \leq \xi-\omega} K(v) \cdot \int_{0}^{\xi} |A(u)|\,du + \underset{u \geq \xi}{\underline{\mathrm{bd}}}\, A(u) \cdot \int_{\xi-\omega}^{\infty} K(v)\,dv \qquad (0 < \xi < \omega),$$

whence we deduce, making $\omega \to \infty$ and then $\xi \to \infty$,

$$(8) \qquad \underset{\omega\to\infty}{\underline{\lim}} \int_{0}^{\infty} A(u)K(u-\omega)\,du \geq \underset{u\to\infty}{\underline{\lim}}\, A(u) \cdot \int_{-\infty}^{\infty} K(v)\,dv = \underset{u\to\infty}{\underline{\lim}}\, A(u) \cdot 2\pi k(0).$$

Also, the second integral in (7) is

$$(9) \qquad \begin{aligned} \int_{-\omega}^{\infty} A^*(\omega+v)K(v)\,dv &= \sum_{-N}^{N} \alpha_n e^{i\gamma_n\omega} \int_{-\omega}^{\infty} K(v)e^{i\gamma_n v}dv \\ &= \sum_{-N}^{N} \alpha_n e^{i\gamma_n\omega} 2\pi k(\gamma_n) + o(1) \end{aligned}$$

as $\omega \to \infty$.

The desired inequality

$$\underset{u\to\infty}{\underline{\lim}}\, A(u) \leq \underset{\omega\to\infty}{\underline{\lim}}\, A^*_T(\omega)$$

now follows from (7), (8) and (9), in virtue of the definitions of $k(t) = k_T(t)$ and of $A^*_T(\omega)$.

THEOREM 2. *Suppose that the complex zeros of* $\zeta(s)$ *are all simple and lie on the line* $\sigma = 1/2$, *and let them be* $(1/2) + i\gamma_n$ $(n = \pm 1, \pm 2, \cdots;$ $\gamma_{-n} = -\gamma_n;$ $0 < \gamma_1 < \gamma_2 < \cdots)$. *Let* α_n *be the residue of* $F(s)$ *at* $s = i\gamma_n$, *where* $F(s)$ *denotes either of the functions*

$$\text{(i)}\ \frac{1}{((1/2)+s)\zeta((1/2)+s)}, \qquad \text{(ii)}\ \frac{\zeta(1+2s)}{((1/2)+s)\zeta((1/2)+s)}.$$

Then $\sum_{1}^{\infty} |\alpha_n|$ *is divergent.*

Consider the integral

$$I = \int (T + is)^k F(s)\, ds$$

taken in the positive sense round the rectangle with vertices $-\alpha, \beta, \beta + iT, -\alpha + iT$, indented upwards at the origin in case (ii), where α and β are fixed ($0 < \alpha < 1/4$, $1/2 < \beta < 1$), k is a fixed positive integer,[2] and $T \to \infty$ through a sequence T_m ($m = 1, 2, \cdots$) such that

$$| 1/\zeta[(1/2) + s)] | < T^K \qquad (-1 \leq \sigma \leq 1, t = T = T_m),$$

where K is some positive constant. The existence of such a sequence T_m is well known (see, e. g., Titchmarsh [7], Theorem 18, p. 26).

By Cauchy's theorem of residues,

$$I = 2\pi i \sum_{0 < \gamma_n < T} (T - \gamma_n)^k \alpha_n = 2\pi i S_k(T),$$

say.

Let I_1, I_2, I_3, I_4 be the contributions to I of the four sides of the rectangle, numbered in the positive sense from the lower horizontal side. We have

$$I_1 = O(T^k),$$

$$\begin{aligned} I_2 &= \sum_{n=1}^{\infty} \frac{a_n}{n^{1/2}} \int_{\beta}^{\beta+iT} \frac{(T + is)^k ds}{((1/2) + s) n^s} \qquad [a_n = \mu(n) \text{ or } \lambda(n)] \\ &= \int_{\beta}^{\beta+iT} \frac{T^k ds}{(1/2) + s} + O(T^k) + O(\sum_{n=2}^{\infty} T^k / n^{1/2+\beta} \log n) \\ &= T^k \log T + O(T^k), \end{aligned}$$

by using the inequality $|(T + is)^k - T^k| \leq K_1 | s | T^{k-1}$ in the term $n = 1$, and deforming the path of integration into the broken line β, T, $T + iT$, $\beta + iT$ [or, alternatively, writing $n^{-s} ds = -d(n^{-s}) / \log n$ and integrating by parts] in the terms $n \geq 2$.

Now the Riemann hypothesis (which we are assuming here) implies that, when $t \to \infty$, $\zeta(1 + 2s) = O(t^{\epsilon})$ uniformly for $-\alpha \leq \sigma \leq \beta$, and

$$\frac{1}{\zeta((1/2) + s)} = \frac{2^{(1/2)-s} \pi^{-(1/2)-s} \cos[(1 + 2s)\pi/4] \Gamma((1/2) + s)}{\zeta((1/2) - s)}$$
$$= O(t^{-\alpha+\epsilon}) \qquad (\sigma = -\alpha),$$

where ϵ is an arbitrarily small positive number. (See, e. g., Titchmarsh [7], § 5. 13, (1), (2).) Hence

---

[2] 'Integer' may be replaced by 'number' if we specify that the branch of $(T + is)^k$ which is real and positive for $s = it$, $t < T$, has to be taken.

$$I_3 = O(1 \cdot T^{-1} \cdot T^{K} \cdot T^{\epsilon}) = O(T^{K-1+\epsilon}),$$

$$I_4 = O\left(\int_0^T T^{k}(t+1)^{-1-\alpha+2\epsilon}dt\right) = O(T^{k}).$$

Collecting these resuts, we obtain

$$I = T^{k}\log T + O(T^{k}) + O(T^{K-1+\epsilon}),$$

so that, if $k > K - 1$,

$$2\pi i S_k(T) \sim T^{k}\log T \tag{10}$$

when $T \rightarrow \infty$ through the sequence T_m. This proves the theorem (and indeed more) since

$$|S_k(T)| \leq T^{k} \sum_{0<\gamma_n<T} |\alpha_n|.$$

3. Proof of Theorem A. We have

$$F(s) = \int_0^\infty A(u)e^{-su}du \qquad (\sigma > 1/2),$$

where $F(s)$ has either of the meanings assigned to it in Theorem 2, and $A(u)$ denotes the corresponding one of the functions

$$\text{(i)} \quad M(e^{u})e^{-u/2}, \qquad \text{(ii)} \quad L(e^{u})e^{-u/2}.$$

The conclusion of Theorem A is equivalent to

$$\overline{\lim_{u\to\infty}}\, A(u) = +\infty, \qquad \underline{\lim_{u\to\infty}}\, A(u) = -\infty. \tag{11}$$

Suppose that the first result (for example) is false. Then there is a constant A such that $A - A(u) \geq 0$ for $u \geq 0$, and the relation

$$\int_0^\infty \{A - A(u)\}e^{-su}du = (A/s) - F(s) = G(s)$$

(say), true in the first instance for $\sigma > 1/2$, holds for $\sigma > 0$ by Landau's theorem, since $G(s)$ is regular along the whole of the positive real axis $s > 0$. Hence $G(s)$, and therefore $F(s)$, is regular in the half-plane $\sigma > 0$; and also

$$|G(s)| \leq G(\sigma) = O(\sigma^{-1}) \quad \text{as} \quad \sigma \rightarrow +0,$$

so that $G(s)$, and therefore $F(s)$, can have no multiple pole on $\sigma = 0$. A similar argument with $A + A(u)$ applies if the second result (11) is false. This means (and all this is classical) that the results (11) are certainly true if $\zeta(s)$ has a complex zero off the line $\sigma = 1/2$, or if all the complex zeros lie on this line but include a multiple zero (and in these cases the hypothesis of linear independence of the γ_n is, of course, irrelevant).

We may, therefore, make the assumptions of Theorem 2. Define γ_n and

α_n as in Theorem 2 for $n = \pm 1, \pm 2, \cdots$, and as follows (in the two cases) for $n = 0$:

$$\text{(i)} \qquad \gamma_0 = 0, \qquad \alpha_0 = 0;$$
$$\text{(ii)} \qquad \gamma_0 = 0, \qquad \alpha_0 = 1/\zeta(1/2).$$

Then the conditions of Theorem 1 are satisfied for any given $T > 0$ if we choose N so that $\gamma_N \geq T$; and the conclusions (5) therefore hold.

But, if the γ_n $(0 < \gamma_n < T)$ are linearly independent,

$$\overline{\lim_{u\to\infty}} A^*_T(u) = \alpha_0 \pm 2 \sum_{0<\gamma_n<T} [1 - (\gamma_n/T)] \mid \alpha_n \mid, \tag{12}$$

since by Kronecker's theorem we can find arbitrarily large values of u for which the products $\gamma_n u$ $(0 < \gamma_n < T)$ are simultaneously as near as we please to $-\arg(\pm \alpha_n) \pmod{2\pi}$. But the sum Σ on the right of (12) is

$$\geq \sum_{0<\gamma_n<T/2} \mid \alpha_n \mid /2$$

and is therefore arbitrarily large with T by Theorem 2. Hence

$$\overline{\lim_{u\to\infty}} A(u) \geq \lim_{T\to\infty} \overline{\lim_{u\to\infty}} A^*_T(u) = +\infty,$$

with a similar result for the lower limit.

The argument is not essentially affected if there are a finite number of relations of the type (3), since these will involve a last γ_n, say γ_M, and we can apply Kronecker's theorem to the γ_n in the range $\gamma_M < \gamma_n < T$.

REFERENCES.

1. F. Mertens, "Über eine zahlentheoretische Funktion," *Sitzungsberichte Akad. Wien*, vol. 106, Abt. 2a (1897), pp. 761-830.
2. G. Pólya, "Verschiedene Bemerkungen zur Zahlentheorie," *Jahresbericht der deutschen Math.-Vereinigung*, vol. 28 (1919), pp. 31-40.
3. R. D. von Sterneck, "Empirische Untersuchung über den Verlauf der zahlentheoretischen Funktion $\sigma(n) = \sum_{x=1}^{x=n} \mu(x)$ im Intervalle von 0 bis 150000," *Sitzungsberichte Akad. Wien*, vol. 106, Abt. 2a (1897), pp. 835-1024.
4. ———, "Empirische Untersuchung . . . von 150000 bis 500000," *ibid.*, vol. 110, Abt. 2a (1901), pp. 1053-1102.
5. ———, "Neue empirische Daten über die zahlentheoretische Funktion $\sigma(n)$," *Proc. 5th International Congress of Mathematicians*, vol. 1 (1912), pp. 341-343.
6. T. J. Stieltjes, *Correspondance d'Hermite et de Stieltjes* (Paris, 1905), Lettre 79.
7. E. C. Titchmarsh, *The zeta-function of Riemann*, Cambridge (1930).

ON THE PRIME NUMBER THEOREM.*

By AUREL WINTNER.

A few years ago, Watson succeeded in proving an asymptotic congruence property formulated by Ramanujan, namely, the following theorem: If m and k are fixed positive integers, there are between $n = 1$ and $n = N$ only $o(N)$ integers n for which the sum of the $(2m-1)$-th powers of all divisors of n is not a multiple of k. This implies, in the particular case $2m - 1 = 11$, that Ramanujan's $\tau(n)$ is divisible by 691 except for $o(N)$ values of n.

Actually, Watson has improved on $o(N)$ by logarithmical factors (for arbitrary m, k), and has even proved asymptotic formulae for certain summatory functions on which the o-estimates depend[1]; asymptotic formulae established by the classical methods of the theory of primes (that is, by contour integrations, based on estimates of $\zeta(s)$ in a certain domain containing the line $\sigma = 1$ in its interior). The singularities of the generating Dirichlet series and, correspondingly, the proof of the asymptotic formulae are similar to those occurring in Landau's result,[2] according to which the number of the positive integers which are less than x and are representable as a sum of two squares is asymptotically proportional to $x/(\log x)^{\frac{1}{2}}$. The following considerations, which deal with the general analytical background of asymptotic laws of this type, seem to be warranted in view of Hardy's recent presentation[3] of the results just mentioned.

The asymptotic formula, given by Hardy in his discussion of Landau's problem, is false, the passage from the order of the singularity (at $s = 1$) to the order of the summatory function being erroneous. This can be seen, without any Tauberian argument, from Abelian reasons alone. (Among all the possible orders, Landau's exponent, $\frac{1}{2}$, represents the only case in which Hardy's calculation becomes correct.) However, my quarrel is not with this *lapsus calami* (which does not occur in Watson's paper), but with the method of approach. In fact, Hardy refers (as Watson did) to the parallel problem of

---

* Received November 10, 1941.

[1] G. N. Watson, "Ueber Ramanujansche Kongruenzeigenschaften der Zerfällungszahlen. (I)," *Mathematische Zeitschrift*, vol. 39 (1935), pp. 712-731.

[2] E. Landau, "Ueber die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindestzahl der zu ihrer additiven Zusammenstezung erforderlichen Quadrate," *Archiv der Mathematik und Physik*, ser. 3, vol. 13 (1908), pp. 305-312.

[3] G. H. Hardy, *Ramanujan*, Cambridge University Press, 1940, pp. 61-63 and 107-109.

two squares; a problem presented along the lines of Landau's approach of 1908, requiring, as the prime number theorem did at that time, a definite amount of information on a domain contained in the half-plane $\sigma < 1$. And Hardy commences the presentation of this approach by a remark concerning its analytical interest, which is observed to be due to the fact that the pole of the prime number theorem becomes replaced by an algebraic singularity; the order of infinity, say μ, at $s = 1$ now being $\frac{1}{2}$ instead of 1. But all of this must be felt as rather disturbing, the more so as, after Wiener's work,[4] the prime number theorem itself has nothing to do with what happens beyond the line $\sigma = 1$.

In the sequel, this methodical anomaly will be disposed of by showing that, on the one hand, there is no point in passing through the line $\sigma = 1$ in order to prove the theorem of either Landau or Watson, since a *unified,* and much shorter, approach to all these problems (including the prime number theorem itself) can be obtained by a straight-forward extension of Ikehara's theorem concerning the case $\mu = 1$ of a simple pole; and that, on the other hand, all these problems do not involve the algebraic character of the underlying singularities, since μ could be irrational, and therefore the singularity at $s = 1$ logarithmic, as far as the unified Tauberian approach is concerned. Even the fact that, in the problems of Landau and Watson, μ not only is rational but less than 1 as well, turns out to be without any significance, since nothing happens when μ passes through Ikehara's case, $\mu = 1$.

The necessary modifications of the proof of Ikehara's theorem[5] are so much on the surface that certain details will be given only because the possibility of these modifications appears to be overlooked in the literature. The only complication arising from the passage to the necessary generalization of Ikehara's case is the appearance of an additional factor. In fact, the Fourier transformation belonging to an arbitrary $\mu > 0$ corresponds to the Γ-integral

$$\int_0^\infty e^{-wx}x^{\mu-1}dx;$$

so that Ikehara's case, $\mu - 1 = 0$, of a simple pole might have disguised the availability of the method in the general situation. The latter can be formulated as follows:

---

[4] N. Wiener, *The Fourier Integral and Certain of its Applications,* Cambridge University Press, 1933, pp. 112-137. Actually, the first proof of the prime number theorem avoiding a crossing of the line $\sigma = 1$ (and based on the characteristic application of the Riemann-Lebesgue lemma for Fourier integrals) is due to Hardy and Littlewood, *Quarterly Journal of Mathematics,* vol. 46 (1915), pp. 215-219. However, their proof presupposes at any rate a rough (exponential) estimate of $\zeta(1 + it)$ for large t. But any such estimate depends on the same machinery as the crossing of the line $\sigma = 1$.

[5] Cf. *loc. cit.*[4], pp. 127-130.

If a Dirichlet series

$$f(s) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n s}, \qquad (0 = \lambda_1 < \lambda_2 < \cdots), \tag{1}$$

where

$$a_n \geqq 0, \tag{2}$$

is convergent for $\sigma > 1$ *and such that there exist two constants,* $C \neq 0$ *and* $\mu > 0$, *for which the regular function*

$$f(s) - C/(s-1)^{\mu}, \text{ where } \sigma > 1, \qquad (\arg(\sigma - 1)^{\mu} = 0), \tag{3}$$

possesses a continuous boundary function as $\sigma \rightarrow 1 + 0$, *then*

$$\sum_{\lambda_n \leqq x} a_n \sim \frac{C}{\Gamma(\mu)} e^{x} x^{\mu - 1} \qquad \text{as } x \rightarrow \infty. \tag{4}$$

It is understood that by the existence of a continuous boundary function is meant this: For every fixed $T > 0$, the function (3) of $s = \sigma + it$ tends, as $\sigma \rightarrow 1 + 0$, to a function of t uniformly for $-T \leqq t \leqq T$. Actually, it is more than sufficient to assume the existence of two functions, say

$$g_0(t) \text{ and } g^*(t) \geqq 0, \qquad (-\infty < t < \infty), \tag{5}$$

such that the function (3) of $s = \sigma + it$ tends, as $\sigma \rightarrow 1 + 0$, to $g_0(t)$ if t is arbitrarily fixed on the complement of a t-set of measure 0, and $g^*(t)$ is L-integrable on every interval $-T \leqq t \leqq T$ and exceeds the absolute value of the difference (3) whenever $1 < \sigma < 2$ (say).

Even without this generalization of the italicized Tauberian theorem, the prime number theorem and Landau's result follow by choosing

$$\mu = 1, \qquad f(s) = -\frac{\zeta'(s)}{\zeta(s)} \left(\text{or, equivalently, } f(s) = \sum_{p} \frac{\log p}{p^{s}} \right)$$[6]

---

[6] The summatory functions (4) belonging to these two ordinary Dirichlet series (the second of which defines a function meromorphic in the half-plane $\sigma > 0$, with poles clustering at every point of the line $\sigma = 0$) are $\psi(\exp x)$ and $\theta(\exp x)$, if the least common multiple of all integers between 1 and x and of all primes between 1 and x are respectively denoted by $\exp \psi(x)$ and $\exp \theta(x)$. It is interesting that the prime number theorem, $\psi(x) \sim x$, is equivalent not only to $\theta(x) \sim x$ but also to an asymptotic formula, $\psi(x) - \theta(x) \sim x^{1/2}$, for the excess represented by multiple prime factors (counted in ψ but not in θ); in fact,

$$\psi(x) = \sum_{n=1}^{\infty} \theta(x^{1/n}), \text{ where } \sum_{n=3}^{\infty} \theta(x^{1/n}) = x^{1/3} O(x^{\epsilon}) = o(x^{1/2}).$$

and

$$\mu = \tfrac{1}{2}, \qquad \{f(s)\}^2 = \zeta(s)\,\frac{L(s)}{1-2^{-s}} \prod_{p\equiv 3 \pmod 4} \left(1 - \frac{1}{p^{2s}}\right)^{-1}$$

respectively, where the L-series is that belonging to the non-principal character (mod 4). Correspondingly, Watson's theorem results by placing

$$\mu = 1 - \frac{1}{\phi(k)}, \qquad f(s) = \zeta(s) \prod_{p\equiv -1 \pmod k} \left(1 + \frac{1}{p^s}\right)^{-1},$$

where ϕ is Euler's function and $k > 2$. Needless to say, the various extensions of these results (for instance, the asymptotic formula for the number of integers represented by a binary quadratic form,[7] where, as in Landau's particular case, $\mu = \tfrac{1}{2}$) follow in the same way.

In order to facilitate a direct comparison with the standard treatment[8] of Ikehara's case, the weight function of Tauberian averaging will be chosen in the usual way, that is, such that the kernel of the Fourier transformation becomes

$$(6) \qquad S(u) = (\sin u/u)^2.$$

Since $|u|^{\mu} S(u) \to 0$ as $|u| \to \infty$ then holds only if

$$(7) \qquad \mu < 2, \qquad (\mu > 0),$$

the choice (6) will necessitate the restriction (7). In order to include arbitrarily large values of μ, all that is necessary is a replacement of the weight function

$$\omega(x) = \mathrm{Max}(0, 1 - |x|), \qquad (-\infty < x < \infty),$$

which leads to (6), by another weight function, leading to a Fourier transform which tends to 0 stronger than $|u|^{-\mu}$ (and is, as (6), non-negative). Such weight functions can be obtained by the well-known process of successive convolutions, for instance.[9] Incidentally, not only (7) but even $\mu \leqq 1$ is satisfied in all the applications referred to above.

---

[7] Cf. P. Bernays, *Ueber die Darstellung von positiven, ganzen Zahlen durch die primitiven, binären quadratischen Formen einer nicht-quadratischen Diskriminante,* Dissertation, Göttingen, 1912.

[8] Cf. E. Landau, "Ueber Dirichletsche Reihen," *Göttinger Nachrichten*, 1932, pp. 525-527.

[9] Cf., e. g., J. Karamata, "Weiterführung der N. Wienerschen Methode," *Mathematische Zeitschrift*, vol. 38 (1934), pp. 701-708 where, for other reasons, the square of (6) is applied (and, therefore, the first auto-convolution of the above weight function is considered).

The weight function $\exp(-\tfrac{1}{2}x^2)$ would have the advantage that the device of successive convolutions becomes unnecessary. But this self-reciprocal function has its own disadvantages, its order at infinity being too low.

If, for $-\infty < x < \infty$,

$$\alpha(x) = e^{-x} \sum_{\lambda_n \leqq x} a_n, \qquad (a_n \geqq 0;\ 0 = \lambda_1 < \lambda_2 < \cdots), \tag{8}$$

then, by (2),

$$e^x \alpha(x) \leqq e^y \alpha(y) \text{ whenever } x < y, \tag{9}$$

and

$$\alpha(x) = 0 \text{ for } -\infty < x < 0. \tag{10}$$

According to (8), the assertion (4) can be written in the form

$$\alpha(x) \sim x^{\mu-1}, \qquad (x \to \infty), \tag{11}$$

if, without loss of generality, $C = \Gamma(\mu)$ in (3).

Instead of (11), it will be sufficient to prove that

$$\int_{-\infty}^{Tx} \alpha(x - v/T) S(v)\, dv \sim \pi x^{\mu-1} \text{ as } x \to \infty, \tag{12}$$

if $T > 0$ is arbitrarily fixed. In fact, since

$$\int_{-\infty}^{\infty} S(v)\, dv = \pi, \text{ by (6)},$$

the passage from (12) to (11) requires the elimination of the weight factor. But it is clear from (8), (9), (10) that this elimination is supplied precisely by the standard Tauberian argument,[10] if (12) holds for an *arbitrary* value of $T > 0$.

According to (8), (9) and (1),

$$f(s) = \int_0^{\infty} e^{-sx} d\{e^x \alpha(x)\} \equiv s \int_0^{\infty} e^x \alpha(x) e^{-sx} dx, \qquad (\sigma > 1);$$

so that, since

$$\Gamma(\mu)/w^{\mu} \equiv \int_0^{\infty} e^{-wx} x^{\mu-1} dx, \qquad (\Re w > 0, \mu > 0),$$

the difference (3), where $C = \Gamma(\mu)$, can be written in the form

$$s \int_0^{\infty} e^{(1-s)x} \alpha(x)\, dx - \int_0^{\infty} e^{(1-s)x} x^{\mu-1} dx, \qquad (\sigma > 1).$$

---

[10] Cf., e. g., *loc. cit.*[8], 1)-2) on pp. 526-527

Thus the assumptions specified after (5) mean that, if

(13) $$s = \sigma + it, \qquad \sigma = 1 + \epsilon > 1, \qquad (-\infty < t < \infty),$$

and

(14) $$F_\epsilon(t) \equiv \int_0^\infty \{\alpha(x) - x^{\mu-1}\} e^{(1-s)x} dx, \qquad (\epsilon > 0),$$

then there exist for $-\infty < t < \infty$ two functions, say $F_0(t)$ and $F^*(t)$, such that

(15) $$F_\epsilon(t) \rightarrow F_0(t) \text{ as } \epsilon \rightarrow 0 \text{ holds for almost all } t$$

and

(16) $$| F_\epsilon(t) | < F^*(t) \text{ for arbitrary } \epsilon \text{ and } t,$$

where $F^*(t)$ is L-integrable on every finite interval $-T \leqq t \leqq T$.

Since the integral (14) is absolutely convergent for all values of s mentioned under (13), it follows by a legitimate interchange of the order of integrations, that

$$\int_{-2T}^{2T} e^{ixt}(1 - \tfrac{1}{2} | t | /T) F_\epsilon(t) dt = 2T \int_0^\infty \{\alpha(u) - u^{\mu-1}\} e^{-\epsilon u} S(Tx - Tu) du$$

is, by virtue of the definition (14) and (6), an identity in the three parameters ϵ, T, x. But (15) and (16) assure that, if $\epsilon \rightarrow 0$, the integral on the left tends to the (finite) limit

$$\int_{-2T}^{2T} e^{ixt}(1 - \tfrac{1}{2} | t | /T) F_0(t) dt.$$

And the integral analogue of the Riemann-Lebesgue lemma shows that the last integral tends to 0 as $x \rightarrow \infty$, if T is fixed. Consequently, there exists a (finite) limit

(17) $$\lim_{\epsilon \rightarrow 0} T \int_0^\infty \{\alpha(u) - u^{\mu-1}\} e^{-\epsilon u} S(Tx - Tu) du,$$

where T and x are arbitrary. And (17) tends to 0 as $x \rightarrow \infty$, if T is arbitrarily fixed.

Accordingly, if u in (17) is replaced by the integration variable $v = Tx - Tu$,

(18) $$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{Tx} \{\alpha(x - v/T) - (x - v/T)^{\mu-1}\} e^{-\epsilon(x-vT)} S(v) dv \rightarrow 0 \text{ as } x \rightarrow \infty,$$

where T (>0) is arbitrary. Furthermore, from (6) and (7),

$$\lim_{\epsilon\to 0}\int_{-\infty}^{Tx}(x-v/T)^{\mu-1}e^{-\epsilon(x-v/T)}S(v)\,dv=\int_{-\infty}^{Tx}(x-v/T)^{\mu-1}S(v)\,dv,$$

and it is also clear from (6) and (7) that the ratio of the last integral to $x^{\mu-1}$ is

$$\equiv\int_{-\infty}^{Tx}\left(1-\frac{v}{Tx}\right)^{\mu-1}\frac{\sin^2 v}{v^2}\,dv\to\int_{-\infty}^{\infty}\frac{\sin^2 v}{v^2}\,dv=\pi,\qquad x\to\infty,$$

where T is arbitrarily fixed. It follows therefore from (18) that

$$(19)\qquad x^{1-\mu}\lim_{\epsilon\to 0}\int_{-\infty}^{Tx}\alpha(x-v/T)e^{-\epsilon(x-v/T)}S(v)\,dv\to\pi,\qquad x\to\infty,$$

holds for every T.

In order to complete the proof of (12), it remains to ascertain that, if both x and T are fixed, the limit process $\epsilon\to 0$ can be carried out beneath the integral sign in (19):

$$(20)\qquad \lim_{\epsilon\to 0}\int_{-\infty}^{Tx}\alpha(x-v/T)e^{-\epsilon(x-v/T)}S(v)\,dv=\int_{-\infty}^{Tx}\alpha(x-v/T)S(v)\,dv.$$

But the existence of the (finite) limit on the left of (20) was proved above. Furthermore, both functions (6), (8) are real and non-negative. Hence, (20) can be reduced to the trivial fact that, for every positive R $(<\infty)$,

$$\lim_{\epsilon\to 0}\int_{-R}^{Tx}\alpha(x-v/T)e^{-\epsilon(x-v/T)}S(v)\,dv=\int_{-R}^{Tx}\alpha(x-v/T)S(v)\,dv.$$

THE JOHNS HOPKINS UNIVERSITY.

MATRIX REPRESENTATIONS OF COMPLETELY SIMPLE SEMIGROUPS.*

By A. H. Clifford.

By a semigroup is meant a system S of elements $a, b, \cdots$ closed under a single binary associative operation:

$$(ab)c = a(bc).$$

To each element a of S let there correspond a uniquely determined matrix $T(a)$ with n rows and columns and with elements in a (commutative) field Ω. If, for all a, b in S,

$$T(ab) = T(a)T(b),$$

then the correspondence $\mathfrak{T}: a \rightarrow T(a)$ is called a (matrix) representation of S in Ω of degree n. The notions of equivalence, reduction, and decomposition are defined exactly as in the theory of representations of groups or algebras.

The only work dealing with representations of semigroups, of which the author is aware, is that of Suschkewitsch. In [2] he makes considerable progress in the determination of all representations of a type of finite semigroup which he calls a Kerngruppe. [3] gives an abstraction of the process used in [2], at the same time removing the finiteness restriction. In a previous paper [1], he determined the structure of all possible Kerngruppen.

The latter determination has recently been extended and simplified by Rees [4]. In Rees's terminology, a Kerngruppe is a completely simple (finite) semigroup without a zero element. He characterizes every completely simple semigroup as a regular matrix semigroup over a group with zero (definitions in § 2). In the present paper we show how to construct all representations of a (not necessarily regular) matrix semigroup over a group with zero; the process is summarized at the close of § 6. In a final section (§ 8) we show how a Brandt groupoid [5] can be made into a matrix semigroup of especially simple type by the adjunction of a zero element; the theory of the present paper is then applied to find all its representations.

1. Factorizations of a matrix of finite rank. This first section is concerned with a problem in pure matrix theory, namely that of finding all solutions X, Y of the matric equation

$$XY = H,$$

---

* Received March 18, 1941.

where H is a given matrix, possibly infinite but of finite rank, X to have a finite number of columns, and Y a finite number of rows.

Let J and K be arbitrary sets, and Ω a field. By a $J \times K$ matrix in Ω we mean any function $A = a(\iota, \kappa) = (a_{\iota\kappa})$ on the product-set $J \times K$ to Ω (ι ranging over J, κ over K). If J is a finite set of m elements we shall also call A an $m \times K$ matrix, with a similar interpretation for a $J \times n$ matrix and an $m \times n$ matrix. By a submatrix of a $J \times K$ matrix $(a_{\iota\kappa})$ we mean a $J' \times K'$ matrix $(a_{\iota'\kappa'})$, where ι' and κ' range over subsets J' and K' of J and K. By the *rank* of A we mean that of the largest finite non-singular square submatrix of A, if such exist, and ∞ otherwise.

Let $\mathfrak{J}$ and $\mathfrak{K}$ denote the linear spaces over Ω consisting respectively of all J-vectors (a_ι) and all K-vectors (a_κ). The subspace $\mathfrak{R}_0$ of $\mathfrak{K}$ spanned by the rows (fixed ι) of A is the *row-space* of A; the subspace $\mathfrak{Q}_0$ of $\mathfrak{J}$ spanned by the columns (fixed κ) of A is the *column-space* of A. The *row-rank* of A is the dimension of $\mathfrak{R}_0$, the *column-rank* that of $\mathfrak{Q}_0$. By an easy extension of the method of proof for the finite case, one readily sees that *the rank, row-rank, and column-rank of any matrix are all equal.* That is, either they are all finite and equal, or all three are infinite.

THEOREM 1.1. *A $J \times K$ matrix H can be expressed as a product*

$$H = QR$$

of a $J \times t$ matrix Q and a $t \times K$ matrix R if and only if it has finite rank $h \leqq t$.

Proof. If $H = QR$ then the rows of H are linear combinations of those of R, which are t in number, and hence the dimension h of the row-space of H cannot exceed t. Conversely, if $h \leq t$, we may take for R any t rows of H of which all other rows are linear combinations. The ι-th row of Q may then be taken to be the coefficients in any expression of the ι-th row of H as a linear combination of those of R.

Two factorizations

$$H = QR = Q'R'$$

of the matrix H, as in Theorem 1.1, will be called *equivalent* if there exists a non-singular $t \times t$ matrix C such that

$$Q' = QC^{-1}, \quad R' = CR.$$

A factorization with $t = h$ will be called *basic.*

THEOREM 1.2. *Any two basic factorizations of a matrix of finite rank are equivalent.*

Proof. The h rows of R are a basis of the row-space $\mathfrak{R}_0$ of H. Each row of H is uniquely expressible as a linear combination of those of R, and hence Q is uniquely determined by R. Since the h rows of R' are also a basis of $\mathfrak{R}_0$ they are a non-singular linear combination of those of R. Hence $R' = CR$, with C a non-singular $h \times h$ matrix. From $H = QR = Q'CR$ and the unicity of Q, follows $Q'C = Q$.

Let H be a $J \times K$ matrix of finite rank h with row-space $\mathfrak{R}_0$ and column-space $\mathfrak{Q}_0$, and let $H = Q_0R_0$ be any basic factorization of H. Let $H = QR$ be any factorization of H into a $J \times t$ matrix Q of rank q and a $t \times K$ matrix R of rank r. Let $\mathfrak{Q}$ be the column-space of Q, and $\mathfrak{R}$ the row-space of R; then $\mathfrak{Q} \supseteq \mathfrak{Q}_0$ and $\mathfrak{R} \supseteq \mathfrak{R}_0$. Let Q_1 be a $J \times (q - h)$ matrix such that the q columns of the $J \times q$ matrix $(Q_0 \quad Q_1)$ form a basis of $\mathfrak{Q}$, and let R_1 be an $(r - h) \times K$ matrix such that the r rows of the $r \times K$ matrix $\begin{pmatrix} R_0 \\ R_1 \end{pmatrix}$ form a basis of $\mathfrak{R}$.

THEOREM 1.3. *The factorization* $H = QR$ *of* H *is equivalent to the factorization*

$$H = (Q_0 \, 0 \, Q_1 \, 0) \begin{pmatrix} R_0 \\ R_1 \\ 0 \\ 0 \end{pmatrix}. \tag{1.1}$$

Proof. By elementary row transformations on R we can find a non-singular $t \times t$ matrix C_1 such that

$$C_1R = \begin{pmatrix} R^* \\ 0 \end{pmatrix}$$

where R^* is an $r \times K$ matrix whose rows form any basis we like of the row-space $\mathfrak{R}$ of R, in particular

$$R^* = \begin{pmatrix} R_0 \\ R_1 \end{pmatrix}.$$

Let

$$QC_1^{-1} = (Q^* \quad A).$$

Then

$$H = QC_1^{-1} \cdot C_1R = Q^*R^*.$$

Since $\mathfrak{R} \supseteq \mathfrak{R}_0$, the rows of H are vectors in $\mathfrak{R}$, and so are unique linear combinations of the rows of R^*. From

$$(Q_0 \, 0) \begin{pmatrix} R_0 \\ R_1 \end{pmatrix} = Q_0R_0 = H,$$

and the unicity of Q^*, it follows that

$$Q^* = (Q_0\, 0).$$

Let

$$C_2 = \begin{pmatrix} I & U \\ 0 & V \end{pmatrix}$$

where I is the $r \times r$ identity matrix, and V is a non-singular $(t-r) \times (t-r)$ matrix. Then

$$C_2 C_1 R = \begin{pmatrix} R^* \\ 0 \end{pmatrix} = C_1 R,$$

while

$$Q C_1^{-1} C_2 = (Q^* \quad B).$$

where

$$B = Q^* U + AV.$$

Hence A may be replaced by any matrix B obtained from A by elementary column transformations on $A\,(U = 0)$ or by adding to any column of A any linear combination of columns of Q^*. But these are sufficient to reduce A to the form $(Q_1\, 0)$, where the columns of Q_1 form a supplement to those of Q_0 to make a basis of $\mathfrak{Q}$. Taking $C = C_2^{-1} C_1$ with this determination of C_2, we have

$$QC^{-1} = QC_1^{-1}C_2 = (Q^*\, Q_1\, 0) = (Q_0\, 0\, Q_1\, 0),$$

$$CR = C_2^{-1} C_1 R = C_1 R = \begin{pmatrix} R^* \\ 0 \end{pmatrix} = \begin{pmatrix} R_0 \\ R_1 \\ 0 \end{pmatrix} = \begin{bmatrix} R_0 \\ R_1 \\ 0 \\ 0 \end{bmatrix}.$$

COROLLARY 1. 31. *Two factorizations $H = QR = Q'R'$ are equivalent if and only if Q and Q' have the same column-space, R and R' the same row-space.*

Proof. Since (1. 1) depends only on the spaces $\mathfrak{Q}$ and $\mathfrak{R}$, both factorizations are equivalent to (1. 1) and hence to each other.

COROLLARY 1. 32. *Beside the evident inequalities $t \geq q \geq h$, $t \geq r \geq h$, we have $t \geq q + r - h$.*

Proof. By counting columns in the $J \times t$ matrix $(Q_0\, 0\, Q_1\, 0)$ we have $t \geq h + (r - h) + (q - h)$.

The inequalities of Corollary 1. 32 can also be expressed in the form

$$t - h \geq t - q \geq 0, \qquad t - h \geq t - r \geq 0,$$
$$t - h \leq (t - q) + (t - r),$$

which is Sylvester's Law of Nullity in case H, P, Q are all finite $t \times t$ matrices.

2. Completely simple semigroups. Rees [4] defines an *ideal* $\mathfrak{a}$ in a semigroup S to be a subset of S such that aS and Sa are contained in a.* A *zero element* of S is an element 0 such that $a \cdot 0 = 0 \cdot a = 0$ for all a in S. S is simple if it contains no other ideal than S itself and the zero ideal consisting of 0 alone (if S contains a zero element). An element e of S is *idempotent* if $e^2 = e$, and *primitive* if there is no idempotent $f \neq 0, e$ such that $ef = fe = f$. A simple S is *completely simple* if (1) every element of S belongs to at least one set eSf, e and f idempotents, and (2) every idempotent is primitive. Rees's main result is that every completely simple semigroup is isomorphic with what he calls a regular matrix semigroup over a group with zero. This is defined in the following way.

Let G be a group with a zero element 0 adjoined. Let J and K be any two sets of indices; denote the elements of J by $i, j, \cdots$, and those of K by $\kappa, \lambda, \cdots$. Let $P = (p_{\kappa i})$ be a fixed $K \times J$ matrix in G, the $p_{\kappa i}$ being arbitrary but fixed elements of G. For $a \neq 0$ in G, let $(a)_{i\kappa}$ denote the $J \times K$ matrix having a in the i-th row and κ-th column, and zeros elsewhere. $(0)_{i\kappa}$ shall mean, for any i and κ, the null $J \times K$ matrix. S shall consist of all the "one-element" matrices $(a)_{i\kappa}$ together with the null matrix. Multiplication in S is defined by

$$(2.1) \qquad (a)_{i\kappa}(b)_{j\lambda} = (ap_{\kappa j}b)_{i\lambda}.$$

(Thus the product of A and B in S is the matrix product APB; no confusion arises from denoting it simply by AB.) The associative law is readily verified, and so S is a semigroup—*a matrix semigroup over G with defining matrix P.* S is called regular if there is an element $\neq 0$ in each row and each column of P, and S is then completely simple. If no element of P is zero, and S is finite (G, J, K all finite) then the non-zero elements of S form a Suschkewitsch Kerngruppe.

According to Rees's Theorem 2.91 ([4] p. 397) it is possible to change P somewhat without changing S. One such change is to multiply the rows of P on the left by arbitrary elements $r_\kappa \neq 0$ of G, and to multiply its columns on the right by elements $q_i \neq 0$:

$$P \to P^* = RPQ = (r_\kappa p_{\kappa i} q_i) = (p^*_{\kappa i}).$$

This can be seen directly by writing

$$[a]_{i\kappa} = (q_i a r_\kappa)_{i\kappa},$$

whence by (2.1),

$$[a]_{i\kappa}[b]_{j\lambda} = [ap^*_{\kappa j}b]_{i\lambda}.$$

It is thus possible to normalize P so that the elements in a fixed row and a fixed column are either 0 or the identity element e of G. If P is not identically zero we can choose our notation so that, after normalization, $p_{11} = e$ and the elements $p_{\kappa 1}$, $p_{1\iota}$ are either 0 or e.

We shall make explicit use only of $p_{11} = e$. From this and (2.1) we have the following, which we shall need later:

$$(a)_{11}(b)_{11} = (ab)_{11} \tag{2.2}$$

$$(e)_{11}(e)_{\iota 1} = (p_{1\iota})_{11} \tag{2.3}$$

$$(e)_{\iota 1}(e)_{11} = (e)_{\iota 1} \tag{2.4}$$

$$(e)_{11}(e)_{1\kappa} = (e)_{1\kappa} \tag{2.5}$$

$$(e)_{1\kappa}(e)_{11} = (p_{\kappa 1})_{11} \tag{2.6}$$

$$(e)_{1\kappa}(e)_{\iota 1} = (p_{\kappa\iota})_{11} \tag{2.7}$$

$$(e)_{\iota 1}(a)_{11}(e)_{1\kappa} = (a)_{\iota 1}(e)_{1\kappa} = (a)_{\iota\kappa}. \tag{2.8}$$

We shall assume throughout the paper that S is a matrix semigroup over a group G with zero, multiplication being defined by (2.1), *and with the defining matrix P normalized so that $p_{11} = e$.* We do not need to assume that S is regular, but merely that P is not identically zero.

3. Construction of all representations of S as extensions of those of G. If $\mathfrak{T}: a \rightarrow T(a)$ is any representation of S, then $T(0)$ is an idempotent matrix and in the usual fashion $\mathfrak{T}$ can be transformed so that

$$T(0) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

From $0a = a0 = 0$ we readily find, for any a in S,

$$T(a) = \begin{pmatrix} I & 0 \\ 0 & T_2(a) \end{pmatrix}.$$

Hence $\mathfrak{T}$ decomposes into an identity representation $\mathfrak{T}_1: T_1(a) = I$, and a representation $\mathfrak{T}_2$ in which $T_2(0) = 0$. Consequently there is no appreciable loss in generality if *we incorporate into the definition of a representation $\mathfrak{T}$ the requirement that*

$$T(0) = 0.$$

By (2.2), the elements $(a)_{11}$ of S form a group with zero G_1 isomorphic with G. Any representation $\mathfrak{T}^*$ of S induces a representation of G_1, and in the usual fashion $\mathfrak{T}^*{}_1$ can be transformed so that

$$T^*[(a)_{11}] = \begin{pmatrix} T(a) & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.1}$$

where $\mathfrak{T}: a \to T(a)$ is a proper representation of G:

$$(3.2)\qquad T(ab) = T(a)T(b), \qquad T(e) = I, \qquad T(0) = 0.$$

We call $\mathfrak{T}^*$ an *extension of* $\mathfrak{T}$ *from* G *to* S. The possibility that $\mathfrak{T}^*$ may induce a null representation in G_1 need not trouble us, since then by (2.8) $\mathfrak{T}^*$ is a null representation of S.

Corresponding to (3.1), let us write

$$T^*[(e)_{11}] = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}.$$

Then from (2.3) and (2.4) we have

$$\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} T(p_{1\iota}) & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R_{11} & 0 \\ R_{21} & 0 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}.$$

Hence $R_{11} = T(p_{1\iota})$, $R_{12} = 0$, $R_{22} = 0$. Denoting R_{21} by R_ι we have

$$(3.3)\qquad T^*[(e)_{\iota 1}] = \begin{pmatrix} T(p_{1\iota}) & 0 \\ R_\iota & 0 \end{pmatrix}.$$

Similarly, from (2.5) and (2.6) we find

$$(3.4)\qquad T^*[(e)_{1\kappa}] = \begin{pmatrix} T(p_{\kappa 1}) & Q_\kappa \\ 0 & 0 \end{pmatrix}.$$

Finally, by (2.8), $T^*[(a)_{\iota\kappa}]$ is the product of the three matrices (3.3), (3.1), and (3.4). Using (3.2),

$$(3.5)\qquad T^*[(a)_{\iota\kappa}] = \begin{pmatrix} T(p_{1\iota}ap_{\kappa 1}) & T(p_{1\iota}a)Q_\kappa \\ R_\iota T(ap_{\kappa 1}) & R_\iota T(a)Q_\kappa \end{pmatrix}.$$

If the degree of $\mathfrak{T}$ is n and that of $\mathfrak{T}^*$ is $n + t$, then the Q_κ are $n \times t$ matrices, and the R_ι are $t \times n$ matrices. We have also

$$(3.6)\qquad Q_1 = 0, \qquad R_1 = 0.$$

THEOREM 3.1. *Let* $\mathfrak{T}$ *be a proper representation of* G. *Then* (3.5) *defines a representation* $\mathfrak{T}^*$ *of* S *if and only if the matrices* Q_κ *and* R_ι *satisfy*

$$(3.7)\qquad Q_\kappa R_\iota = T(p_{\kappa\iota}) - T(p_{\kappa 1}p_{1\iota}).$$

Conversely, every representation of S *is equivalent to one of this form.*

Proof. Taking the converse first, we have already seen that any representation $\mathfrak{T}^*$ of S can be transformed into the form (3.5), and all that remains is to prove (3.7). From (2.7), (3.4), (3.3), and (3.1) we have

$$\begin{pmatrix} T(p_{\kappa 1}) & Q_\kappa \\ 0 & 0 \end{pmatrix}\begin{pmatrix} T(p_{1i}) & 0 \\ R_i & 0 \end{pmatrix} = \begin{pmatrix} T(p_{\kappa i}) & 0 \\ 0 & 0 \end{pmatrix}$$

which gives (3. 7) immediately.

Following Suschkewitsch [2], we write (3. 5) in the form

$$T^*[(a)_{i\kappa}] = \bar{R}_i T(a) \bar{Q}_\kappa$$

where

$$\bar{R}_i = \begin{pmatrix} T(p_{1i}) \\ R_i \end{pmatrix}, \qquad \bar{Q}_\kappa = (T(p_{\kappa 1}) \quad Q_\kappa).$$

(3. 7) becomes

$$\bar{Q}_\kappa \bar{R}_i = T(p_{\kappa i}).$$

Then

$$T^*[(a)_{i\kappa}]T^*[(b)_{j\lambda}] = \bar{R}_i T(a) \bar{Q}_\kappa \bar{R}_j T(b) \bar{Q}_\lambda = \bar{R}_i T(a) T(p_{\kappa j}) T(b) \bar{Q}_\lambda$$
$$= \bar{R}_i T(a p_{\kappa j} b) \bar{Q}_\lambda = T^*[(a p_{\kappa j} b)_{i\lambda}] = T^*[(a)_{i\kappa}(b)_{j\lambda}].$$

We remark that if S has no divisors of zero, and we normalize its defining matrix P so that $p_{1i} = p_{\kappa 1} = e$ for all i and κ, then (3. 5) and (3. 7) simplify to

$$T^*[(a)_{i\kappa}] = \begin{pmatrix} T(a) & T(a)Q_\kappa \\ R_i T(a) & R_i T(a) Q_\kappa \end{pmatrix},$$
$$Q_\kappa R_i = T(p_{\kappa i}) - I.$$

For a given representation $\mathfrak{T}$ of G we let

$$H_{\kappa i} = T(p_{\kappa i}) - T(p_{\kappa 1} p_{1i}) \tag{3.8}$$

and define $H = H(\mathfrak{T})$ to be the matrix of matrices

$$H = \begin{pmatrix} & \vdots & \\ \cdots & H_{\kappa i} & \cdots \\ & \vdots & \end{pmatrix} \qquad (\kappa \neq 1, i \neq 1). \tag{3.9}$$

Then the equations (3. 7) can be expressed as a single matrix equation

$$H = QR \tag{3.10}$$

where

$$Q = \begin{pmatrix} \vdots \\ Q_\kappa \\ \vdots \end{pmatrix}, \qquad R = (\ldots R_i \ldots) \qquad (\kappa \neq 1, i \neq 1).$$

We exclude $\kappa = 1$ and $i = 1$ since H_{1i} and $H_{\kappa 1}$ are zero and we have (3. 6). (While it would not be necessary to adhere to (3. 6) at this point, it is essential that we do so when we discuss equivalence in the next section.) In case J (or K) has only one element there is no need to define H at all; in this case the matrices Q_κ (or R_i) are entirely arbitrary.

From Theorem 1. 1 we have immediately:

THEOREM 3.2. *A given representation $\mathfrak{T}$ of G of degree n has an extension $\mathfrak{T}^*$ to S of finite degree if and only if the rank h of the matrix H defined by* (3.8) *and* (3.9) *is finite. The defining matrices Q_κ, R_i of any extension $\mathfrak{T}^*$ of $\mathfrak{T}$ of degree $n + t$ are obtained from a factorization* (3.10) *of H. This is possible if and only if $t \geq h$, and so every extension of $\mathfrak{T}$ has degree at least $n + h$.*

4. Equivalence of two extensions. Two equivalent representations $\mathfrak{T}^*$ and $\mathfrak{T}'^*$ of S evidently induce equivalent representations $\mathfrak{T}$ and $\mathfrak{T}'$ of G, and hence $\mathfrak{T}'^*$ can be transformed so that $\mathfrak{T}' = \mathfrak{T}$. (3.5) for $\mathfrak{T}'^*$ is then of the form

$$T'^*[(a)_{i\kappa}] = \begin{pmatrix} T(p_{1i}ap_{\kappa 1}) & T(p_{1i}a)Q'_\kappa \\ R'_iT(ap_{\kappa 1}) & R'_iT(a)Q'_\kappa \end{pmatrix},$$

with

$$Q'_1 = 0, \qquad R'_1 = 0.$$

THEOREM 4.1. *The two extensions $\mathfrak{T}^*$ and $\mathfrak{T}'^*$ to S of degree $n + t$ of the same representation $\mathfrak{T}$ of G of degree n are equivalent if and only if*

$$Q' = C_1QC_2^{-1}, \quad i.e. \quad Q'_\kappa = C_1Q_\kappa C_2^{-1} \qquad (all\ \kappa \neq 1),$$
$$R' = C_2RC_1^{-1}, \quad i.e. \quad R'_i = C_2R_iC_1^{-1} \qquad (all\ i \neq 1);$$

where C_1 is a non-singular $n \times n$ matrix commuting with $\mathfrak{T}$,

$$C_1T(a) = T(a)C_1 \qquad (all\ a\ in\ G),$$

and C_2 is a non-singular $t \times t$ matrix.

Proof. Suppose that $\mathfrak{T}^*$ and $\mathfrak{T}'^*$ are equivalent. Then there exists a constant non-singular $(n + t) \times (n + t)$ matrix

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

such that

$$CT^*[(a)_{i\kappa}] = T'^*[(a)_{i\kappa}]C.$$

For $i = 1$, $\kappa = 1$ this gives

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}\begin{pmatrix} T(a) & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} T(a) & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

and since $T(a)$ is non-singular we find

$$C_{12} = 0, \qquad C_{21} = 0, \qquad C_{11}T(a) = T(a)C_{11}.$$

Hence we may write

$$C = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$$

where C_1 is a non-singular $n \times n$ matrix commuting with $\mathfrak{T}$, and C_2 is a non-singular $t \times t$ matrix. Then

$$CT^*[(a)_{\iota\kappa}] = \begin{pmatrix} C_1T(p_{1\iota}ap_{\kappa 1}) & C_1T(p_{1\iota}a)Q_\kappa \\ C_2R_\iota T(ap_{\kappa 1}) & C_2R_\iota T(a)Q_\kappa \end{pmatrix},$$

$$T'^*[(a)_{\iota\kappa}]C = \begin{pmatrix} T(p_{1\iota}ap_{\kappa 1})C_1 & T(p_{1\iota}a)Q'_\kappa C_2 \\ R'_\iota T(ap_{\kappa 1})C_1 & R'_\iota T(a)Q'_\kappa C_2 \end{pmatrix}.$$

For $a = e$, $\iota = 1$, equating these gives $C_1Q_\kappa = Q'_\kappa C_2$. For $a = e$, $\kappa = 1$, we find $C_2R_\iota = R'_\iota C_1$. In terms of the composite matrices introduced in (3.10), these may be written

$$C_1Q = Q'C_2, \qquad C_2R = R'C_1.$$

But conversely, if these equations hold, and C_1 commutes with all $T(a)$, then the two matrices above are equal. To verify the southeast corner:

$$C_2R_\iota T(a)Q_\kappa = R'_\iota C_1T(a)Q_\kappa = R'_\iota T(a)C_1Q_\kappa = R'_\iota T(a)Q'_\kappa C_2.$$

If C is non-singular, then $\mathfrak{T}^*$ and $\mathfrak{T}'^*$ are equivalent.

THEOREM 4.2. *If, in the preceding theorem, $\mathfrak{T}$ is absolutely irreducible, then $\mathfrak{T}^*$ and $\mathfrak{T}'^*$ are equivalent if and only if*

$$Q' = QC^{-1}, \qquad R' = CR,$$

where C is a non-singular $t \times t$ matrix. This in turn is equivalent to the condition that Q and Q' have the same column-space, R and R' the same row-space.

Proof. Since C_1 must now be a scalar matrix $\neq 0$, by Schur's Lemma, we can take $C = C_2C_1^{-1}$. But this asserts that the factorizations $H = QR = Q'R'$ of H are equivalent, as defined in § 1. The second part follows immediately from Corollary 1.31.

This result makes it clear that, unlike the situation in group theory, even a finite S will in general have an infinite number of inequivalent representations of the same degree.

5. Proper extensions. A representation $\mathfrak{T}^*$ of S will be called *proper* if it does not decompose into two representations one of which is a null representation.

THEOREM 5.1. *Let $\mathfrak{T}$ be a representation of degree n of G, and let the rank h of the matrix H defined by* (3.8) *and* (3.9) *be finite. Let $\mathfrak{T}^*$ be an extension to S of degree $n + t$ of $\mathfrak{T}$, and let the ranks of its defining matrices Q and R be q and r. Then $t \geq q + r - h$. $\mathfrak{T}^*$ is proper if and only if*

$t = q + r - h$. *If* $t > q + r - h$, *then* $\mathfrak{T}^*$ *decomposes into a proper extension* $\mathfrak{T}'^*$ *of degree* $n + (q + r - h)$ *and a null representation of degree* $t - (q + r - h)$.

Proof. $t \geqq q + r - h$ by Corollary 1.32. Let us assume first that $t > q + r - h$. By Theorem 4.1 with $C_1 = I$, and Theorem 1.3, $\mathfrak{T}^*$ is equivalent to an extension of $\mathfrak{T}$ with defining matrices of the form

$$Q = (Q^0 \ 0 \ Q^1 \ 0), \qquad R = \begin{pmatrix} R^0 \\ R^1 \\ 0 \\ 0 \end{pmatrix}, \tag{5.1}$$

This means of course that

$$Q_\kappa = (Q_\kappa{}^0 \ 0 \ Q_\kappa{}^1 \ 0), \qquad R_\iota = \begin{pmatrix} R_\iota{}^0 \\ R_\iota{}^1 \\ 0 \\ 0 \end{pmatrix}. \tag{5.2}$$

Let us write (5.1) in the form

$$Q = (Q' \ 0), \qquad R = \begin{pmatrix} R' \\ 0 \end{pmatrix} \tag{5.3}$$

where the 0 in Q is a null matrix of $t - (q + r - h)$ columns, and the 0 in R has the same number of rows. Then

$$Q_\kappa = (Q'_\kappa \ 0), \qquad R_\iota = \begin{pmatrix} R'_\iota \\ 0 \end{pmatrix},$$

and putting these in (3.5) we find immediately

$$T^*[(a)_{\iota\kappa}] = \begin{pmatrix} T'^*[(a)_{\iota\kappa}] & 0 \\ 0 & 0 \end{pmatrix} \tag{5.4}$$

where $\mathfrak{T}'^*$ has defining matrices Q' and R'.

We have thus shown that if $t > q + r - h$ then $\mathfrak{T}^*$ is not proper, but decomposes into $\mathfrak{T}'^*$ of degree $n + (q + r - h)$ and a null representation of degree $t - (q + r - h)$. The ranks of Q' and R' are evidently q and r. Hence all that remains to complete the proof of the theorem is to show that if $t = q + r - h$ then $\mathfrak{T}^*$ is proper.

But if $\mathfrak{T}^*$ were not proper, it could be put in the form (5.4) with proper $\mathfrak{T}'^*$. The degree of $\mathfrak{T}'^*$ would then be $n + t'$, where $t' = q' + r' - h$, q' and r' the ranks of the defining matrices Q' and R' of $\mathfrak{T}'^*$. From the form of (5.4) it follows directly that Q and R have the form (5.3). Hence $q = q'$, $r = r'$. But this means $t = t'$, $\mathfrak{T}^* = \mathfrak{T}'^*$.

6. The basic extension; normal forms. Let $\mathfrak{T}$ be a representation of G such that the matrix H defined by (3. 8) and (3. 9) has finite rank h. Let

$$H = Q^0 R^0$$

be a basic factorization of H (§ 1), the number of columns of Q^0 and the number of rows of R^0 being h. By Theorem 1. 2 any two such factorizations are equivalent, and hence the following is an immediate consequence of Theorem 4. 1.

THEOREM 6. 1. *A representation $\mathfrak{T}$ of G of degree n, with H of finite rank h, possesses to within equivalence exactly one extension $\mathfrak{T}^*_0$ to S of degree $n + h$.*

We call $\mathfrak{T}_0$ the *basic* extension of $\mathfrak{T}$ to S.

THEOREM 6. 2. *Every extension $\mathfrak{T}^*$ to S of a representation $\mathfrak{T}$ of G can be transformed into the form*

$$(6.\,2)\quad T^*[(a)_{\iota\kappa}] = \begin{pmatrix} 0 & R_\iota^1 T(ap_{\kappa 1}) & R_\iota^1 T(a) Q_\kappa^0 & R_\iota^1 T(a) Q_\kappa^1 \\ 0 & T(p_{1\iota} a p_{\kappa 1}) & T(p_{1\iota} a) Q_\kappa^0 & T(p_{1\iota} a) Q_\kappa^1 \\ 0 & R_\iota^0 T(ap_{\kappa 1}) & R_\iota^0 T(a) Q_\kappa^0 & R_\iota^0 T(a) Q_\kappa^1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The diagonal block inside the dotted lines is the basic extension $\mathfrak{T}^*_0$ of $\mathfrak{T}$. $\mathfrak{T}^*$ is proper if and only if the rank of the composite matrix Q^1 is equal to the number of its columns, and that of R^1 to the number of its rows.

Proof. As in the proof of Theorem 5. 1, $\mathfrak{T}^*$ can be transformed so that its defining matrices Q and R have the form (5. 1), where $H = Q^0 R^0$ is a basic factorization of H, Q^1 is a matrix of rank $q - h$ equal to the number of its columns, and R^1 is a matrix of rank $r - h$ equal to the number of its rows. $\mathfrak{T}^*$ is proper if and only if the tail null-matrices in Q and R are absent, and then (5. 2) becomes

$$Q_\kappa = (Q_\kappa^0 \;\; 0 \;\; Q_\kappa^1), \qquad R_\iota = \begin{pmatrix} R_\iota^0 \\ R_\iota^1 \\ 0 \end{pmatrix}.$$

Substituting these in (3. 5), and then permuting cyclically the first three rows and columns of the result, we obtain (6. 2).

An improper $\mathfrak{T}^*$ can be put in the form (6. 2) bordered with null matrices to the right and below. This is also of the form (6. 2) with Q^1 replaced by $(Q^1 \;\; 0)$, and the latter has rank less than the number of its columns.

This result makes the construction of all representations of S as direct as one could reasonably expect, provided of course that we know how to construct all those of G:

(1) Construct all representations of G, and discard all except those $\mathfrak{T}$ for which the matrix H defined by (3.8) and (3.9) has finite rank h.

(2) For given $\mathfrak{T}$, construct a basic factorization $H = Q^0R^0$ of H. To do this one may take for R^0 any h linearly independent rows of H; Q^0 is then uniquely determined, each row being the coefficients in the expression of the corresponding row of H as linear combinations of the rows of R^0.

(3) With the resulting $n \times h$ matrices $Q_\kappa{}^0$ and $h \times n$ matrices $R_\iota{}^0$, as given by (3.10), write down the basic extension $\mathfrak{T}^*{}_0$ from (3.5).

(4) Every representation $\mathfrak{T}^*$ of S is an extension of some $\mathfrak{T}$, and any such is obtained simply by writing down (6.2), the matrices $Q_\kappa{}^1$ and $R_\iota{}^1$ being entirely arbitrary. $\mathfrak{T}^*$ will be proper (§ 5) if and only if the composite matrices Q^1 and R^1 have maximum possible rank.

7. Reduction and decomposition. In spite of having an explicit form (6.2) for the representations of S, it is by no means easy to answer all possible questions concerning their reduction and decomposition. We give in this section a few simple results in this direction.

It is clear from the form of (6.2) that a reduction (not a decomposition) of the basic $\mathfrak{T}^*{}_0$ yields a reduction of $\mathfrak{T}^*$, since $\mathfrak{T}^*$ reduces into $\mathfrak{T}^*{}_0$ and a pair of null representations. Thus every irreducible (non-null) constituent of $\mathfrak{T}^*{}_0$ occurs with the same multiplicity in $\mathfrak{T}^*$. But it must be borne in mind that results like this do not tell the whole story when—as in the present case—full reducibility does not hold.

THEOREM 7.1. *If $\mathfrak{T}$ is an irreducible representation of G, then its basic extension $\mathfrak{T}^*{}_0$ to S is also irreducible. Every proper extension $\mathfrak{T}^*$ of $\mathfrak{T}$ is indecomposable, but reduces into $\mathfrak{T}^*{}_0$ and null representations.*

Proof. Were $\mathfrak{T}^*{}_0$ reducible, there would exist a non-singular $(n + h) \times (n + h)$ matrix C such that

$$CT^*{}_0[(a)_{\iota\kappa}]C^{-1} = \begin{pmatrix} T^*{}_1[(a)_{\iota\kappa}] & T^*{}_{12}[(a)_{\iota\kappa}] \\ 0 & T^*{}_2[(a)_{\iota\kappa}] \end{pmatrix}.$$

Here n is as usual the degree of $\mathfrak{T}$, $n + h$ that of $\mathfrak{T}^*{}_0$. But then either $\mathfrak{T}^*{}_1$ or $\mathfrak{T}^*{}_2$ must be an extension to S of the irreducible representation $\mathfrak{T}$ of G. Since they both have degree less than $n + h$, this is impossible by Theorem 3.2.

If $\mathfrak{T}^*$ is any extension of $\mathfrak{T}$, and it decomposes into $\mathfrak{T}^*{}_1$ and $\mathfrak{T}^*{}_2$, then only one of the latter can be an extension to S of the irreducible $\mathfrak{T}$. The other must therefore be a null representation, and $\mathfrak{T}^*$ is not proper. The reduction of $\mathfrak{T}^*$ into $\mathfrak{T}^*{}_0$ and null representations is clear from (6.2).

THEOREM 7.2. *If the representation $\mathfrak{T}$ of G decomposes into $\mathfrak{T}'$ and $\mathfrak{T}''$ then the basic extension $\mathfrak{T}^*_0$ of $\mathfrak{T}$ to S decomposes into the basic extensions $\mathfrak{T}'^*_0$ and $\mathfrak{T}''^*_0$ of $\mathfrak{T}'$ and $\mathfrak{T}''$ respectively.*

Proof. Writing

$$T(a) = \begin{pmatrix} T'(a) & 0 \\ 0 & T''(a) \end{pmatrix}$$

we have

$$H_{\kappa i} = \begin{pmatrix} H'_{\kappa i} & 0 \\ 0 & H''_{\kappa i} \end{pmatrix}.$$

From this it is clear that if any two of the matrices H, H', H'' have finite rank, then so does the third, and $h = h' + h''$. Let n' and n'' be the degrees of $\mathfrak{T}'$ and $\mathfrak{T}''$; that of $\mathfrak{T}$ is then $n = n' + n''$. $\mathfrak{T}'^*_0$ and $\mathfrak{T}''^*_0$ have degrees $n' + h'$ and $n'' + h''$ respectively, by Theorem 6.1. But then their direct sum $\mathfrak{T}'^*_0 + \mathfrak{T}''^*_0$ is an extension to S of $\mathfrak{T}' + \mathfrak{T}'' = \mathfrak{T}$ of degree

$$(n' + h') + (n'' + h'') = n + h,$$

and so must be equivalent to the basic extension $\mathfrak{T}^*_0$ of $\mathfrak{T}$.

Anyone interested will find it worth his while to check this theorem (and the methods of the present paper) on the example given by Suschkewitsch at the close of [2].

8. Application to Brandt groupoids. Let S^* be a groupoid, as defined by Brandt [5]. Adjoin to S^* a zero element 0, and define $ab = 0$ if ab is undefined in S^*. The resulting system S is then a semigroup satisfying (in addition to closure and associativity) the conditions:

I. *If $abc = 0$ then $ab = 0$ or $bc = 0$.*

II. *Each element $a \neq 0$ of S has a unique left identity e and right identity f: $ea = af = a$.*

III. *The equation $ax = b$ [$ya = b$], with given a and b both $\neq 0$, has a unique solution x [y] if a and b have the same left [right] identity.*

IV. *If e and f are idempotents, the set eSf contains an element $\neq 0$.*

Conversely, the non-zero elements of such an S form a Brandt groupoid. We shall therefore call S a *groupoid with zero*. (I is a consequence of II and III; we retain it only for purposes of comparison with Brandt's axioms.)

We show first that S is simple. Let a and b be arbitrary non-zero elements of S; we are to show that the equation $yax = b$ is solvable. Let e be the left identity of a, and f the right identity of b, as given by II. From their unicity follows $e^2 = e$, $f^2 = f$. By IV there exists $c \neq 0$ such that $ec = cf = c$. By III we can solve $ax = c$ and $yc = b$, whence $yax = b$.

By II every element of S belongs to at least (in fact exactly) one set eSf, with e and f idempotent. Because of the unicity in II, every idempotent is primitive. Hence S is completely simple (§ 2). According to Rees's Theorem 2. 93, S can be represented as a regular matrix semigroup over a group G with zero. Let multiplication in S be given by (2. 1).

Let i be a fixed index in J. Since S is regular there exists at least one κ in K such that $p_{\kappa i} \neq 0$. If $p_{\lambda i} \neq 0$ also, then

$$(p_{\kappa i}^{-1})_{i\kappa}(p_{\lambda i}^{-1})_{i\lambda} = (p_{\lambda i}^{-1})_{i\lambda} = (p_{\lambda i}^{-1})_{i\lambda}(p^{-1}{}_{\lambda i})_{i\lambda}.$$

II then requires that $\lambda = \kappa$. Similarly, to a given κ in K there exists exactly one i in J such that $p_{\kappa i} \neq 0$. Hence J and K are in one-to-one correspondence. We may therefore assume that $J = K$, and by suitable numeration choose p_{ii} to be the non-zero element in the i-th row or column of P. As described in § 2, P can then be normalized so that each $p_{ii} = e$. Thus we have the following theorem, the converse being readily verified.

THEOREM 8. 1. *Any Brandt groupoid with zero can be represented as a regular matrix semigroup S over a group G with zero such that $J = K$ and the defining matrix P is the identity matrix:*

$$P_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ e & \text{if } i = k. \end{cases}$$

Conversely, any such S is a Brandt groupoid with zero.

Thus (2. 1) becomes ordinary matrix multiplication:

$$(a)_{ij}(b)_{kl} = \begin{cases} (ab)_{il} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

We proceed now to obtain all representations of S. For i and j both $\neq 1$, (3. 8) becomes

$$H_{ij} = T(p_{ij}),$$

and hence, for any $\mathfrak{X}$, H of (3. 9) is the identity matrix. By Theorem 3. 2, S will have non-null representations of finite degree if and only if J has a finite number s of elements.

Assuming this to be the case, H is the identity matrix of degree $n(s-1)$, if n is the degree of $\mathfrak{X}$. Thus its rank is

$$h = n(s-1),$$

and a basic factorization of H is obtained by taking

$$Q^0 = R^0 = I_h \; (= H).$$

Then the $n \times h$ matrix Q_i^0 can be described as a row of square matrices of degree n, one of them being the identity matrix I and the rest zero; I occurs at the i-th place if we number them from 2 to s. The $h \times n$ matrix R_i^0 is similarly described as a column. Putting these in (3.5), $T^*_0[(a)_{ij}]$ emerges as a matrix of degree s in matrices of degree n, having $T(a)$ in the ij-position and null matrices elsewhere.

If $\mathfrak{T} = \mathfrak{E}$, the identity representation $E(a) = 1$ of G, we obtain thereby the *fundamental* representation $\mathfrak{E}^*$ of degree s of S in which $(a)_{ij}$ is represented by the matrix E_{ij} having 1 in the ij-position and zeros elsewhere. $\mathfrak{T}^*_0$ is thus the direct product of $\mathfrak{E}^*$ and $\mathfrak{T}$ in the sense that

$$T^*_0[(a)_{ij}] = E_{ij} \times T(a), \tag{8.1}$$

although $(a)_{ij} \to T(a)$ is not itself a representation of S.

Consider now any proper extension $\mathfrak{T}^*$ of $\mathfrak{T}$ of degree $n + t$. Let q and r be the ranks of its defining matrices Q and R. Since Q is an $h \times t$ matrix and R a $t \times h$ matrix, q and r are $\leq h$. From Theorem 5.1 we have

$$t = q + r - h \leqq h + h - h = h.$$

But $t \geq h$ by Theorem 3.2. Hence $n + t = n + h$, and $\mathfrak{T}^*$ is the basic extension of $\mathfrak{T}$ by Theorem 6.1.

THEOREM 8.2. *A groupoid S with zero admits non-null representations of finite degree if and only if it has finite degree s over G. In that case, every proper extension $\mathfrak{T}^*$ to S of a representation $\mathfrak{T}$ of G is the basic extension of $\mathfrak{T}$, and is the direct product of $\mathfrak{T}$ with the fundamental representation $\mathfrak{E}^*$ of degree s of S, in the sense of* (8.1).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] A. Suschkewitsch, "Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit," *Mathematische Annalen*, vol. 99 (1928), pp. 30-50.

[2] A. Suschkewitsch, "Über die Matrizendarstellung der verallgemeinerten Gruppen," *Communications de la Société Mathématique de Kharkow*, vol. 6 (1933), pp. 27-38.

[3] A. Suschkewitsch, "Über eine Verallgemeinerung der Semigruppen," *Communications de la Société Mathématique de Kharkow*, vol. 12 (1935), pp. 89-97

[4] D. Rees, "On semi-groups," *Proceedings of the Cambridge Philosophical Society*, vol. 36 (1940), pp. 387-400.

[5] H. Brandt, "Über eine Verallgemeinerung des Gruppenbegriffes," *Mathematische Annalen*, vol. 96 (1927), pp. 360-366.

THE PARAMETRIC OSCULATING QUADRICS OF A FAMILY OF CURVES ON A SURFACE.*

By P. O. BELL.

1. **Introduction.** Dan Sun[1] introduced two interesting quadrics each of which is geometrically determined at a point of a surface by an asymptotic curve of the surface and a one-parameter family of curves on the surface. His definition is essentially the following: At three neighboring points P, P_1, P_2 on an asymptotic curve C_u of a surface S construct the tangents to the curves of a one-parameter family on S. The three tangents determine a quadric whose limit, as P_1, P_2 independently approach P along C_u is the *parametric osculating quadric* $Q^{(u)}$ of the one-parameter family of curves of S at P. The *parametric osculating quadric* $Q^{(v)}$ is similarly defined with respect to the other asymptotic curve C_v.

However, certain errors which Sun made in his routine calculation of the coefficients of the equations of the quadrics $Q^{(u)}$, $Q^{(v)}$[2] invalidate his principal results.

The purpose of this note is to obtain the equations of the quadrics $Q^{(u)}$, $Q^{(v)}$ by a method different from that employed by Sun and to determine some of the properties of the quadrics.

2. **The equations of the quadrics.** Let a non-developable surface S be referred to its asymptotic net as parametric. The homogeneous projective coordinates $y^{(i)}$, $(i = 1, 2, 3, 4)$, of a general point P on S are, then, solutions of a system of differential equations, which by suitable transformation can be reduced to *Wilczynski's canonical form*

$$(2.1) \qquad y_{uu} + 2by_v + fy = 0, \qquad y_{vv} + 2a'y_u + gy = 0.$$

Moreover, the coordinates $y^{(i)}$ satisfy the determinant inequality

$$(2.2) \qquad (y, y_u, y_v, y_{uv}) \neq 0.$$

---

* Received May 5, 1941; Presented to the American Mathematical Society, Feb. 28, 1942.

[1] Dan Sun, "The parametric osculating quadrics of a one-parameter family of curves on a surface," *The Tôhoku Mathematical Journal*, vol. 32 (1930), pp. 81-85.

[2] *Loc. cit.*, equations (5) and (6); compare these equations with our equations (2.9) and (2.10). Sun omitted the details of the calculations whereby he obtained his incorrect equations (5) and (6).

An arbitrary one-parameter family of curves on S is defined by the curvilinear equation

$$dv - \lambda du = 0 \tag{2.3}$$

where λ is an arbitrary function of u, v. Let C_λ denote the curve of the family which passes through P.

The tangents at the points of an asymptotic curve C_u to the curves of the family (2.3) form a ruled surface $R_u^{(\lambda)}$.

The definition of the quadric $Q^{(u)}$ is clearly that of the *osculating quadric* along the generator through P of the ruled surface $R_u^{(\lambda)}$. Hence the *asymptotic regulus* along the generator through P of the ruled surface $R_u^{(\lambda)}$ lies on the quadric $Q^{(u)}$. We use this fact to obtain the equation of $Q^{(u)}$.

The homogeneous coordinates of a general point on the ruled surface $R_u^{(\lambda)}$ are given by the form

$$z = y_u + \lambda y_v + wy. \tag{2.4}$$

Since v is fixed, z is a function of two independent parameters u, w. The differential equation for the asymptotic curves of the ruled surface is therefore

$$L du^2 + 2M du dw = 0, \tag{2.5}$$

where $L = (z_{uu}, z_u, z_w, z)$, $M = (z_{wu}, z_u, z_w, z)$. On differentiating (2.4) and making use of (2.1) the following forms are obtained:

$$\begin{aligned} &z_w = y, \qquad z_u = (\lambda_u - 2b)y_v + \lambda y_{uv} + wy_u - fy, \qquad z_{uw} = y_u, \\ &z_{uu} = (2\lambda_u - 2b)y_{uv} + (4a'b\lambda - f)y_u + (\lambda_{uu} - 2b_u - 2b_v\lambda - f\lambda - 2bw)y_v \\ &\qquad + (2bg\lambda - f_v\lambda - fw - f_u)y. \end{aligned}$$

Substituting these forms in (2.5) we find, by means of elementary determinant theory and the inequality (2.2), that the equation for the curved asymptotics of $R_u^{(\lambda)}$ may be expressed in the form

$$\begin{aligned} dw/du = (&\lambda\lambda_{uu} + 2\lambda\lambda_u w + 6b\lambda_u - 2\lambda_u^2 - 2b_u\lambda - 2b_v\lambda^3 \\ &- 4a'b\lambda^3 - 4b\lambda w - 4b^2)/2\lambda^2. \end{aligned} \tag{2.6}$$

The homogenous coordinates of a general point of the asymptotic regulus along the generator of $R_u^{(\lambda)}$ at P are, therefore, given by

$$x = z + t(z_u + [dw/du]z_w), \tag{2.7}$$

wherein w, t are independent parameters and dw/du is defined by (2.6). In view of (2.2) let the points y, y_u, y_v, y_{uv} be selected as vertices of a local tetrahedron of reference and let the unit point be chosen so that local coordinates of the point $x = x_1 y + x_2 y_u + x_3 y_v + x_4 y_{uv}$ are x_1, x_2, x_3, x_4. These coordinates for x are then readily found, by expressing the right member of (2. 7) as a linear combination of y, y_u, y_v, y_{uv}, to be given by

$$(2.8)\qquad \begin{cases} x_1 = w - ft + t\,dw/du, \\ x_2 = 1 + wt, \\ x_3 = \lambda + (\lambda_u - 2b)t, \\ x_4 = \lambda t, \end{cases}$$

where dw/du is defined by (2. 6). Homogeneous elimination of the parameters w, t from among these equations yields the equation of the quadric $Q^{(u)}$

$$(2.9)\quad 2\lambda^3(x_1x_4 - x_2x_3) + 2\lambda^2x_3^2 + 2(2b - \lambda_u)\lambda x_3x_4 + (4b^2 + 2b_u\lambda + 2b_v\lambda^2 + 2f\lambda^2 + 4a'b\lambda^3 - 6b\lambda_u + 2\lambda_u^2 - \lambda\lambda_{uu})x_4^2 = 0.$$

The equation for $Q^{(v)}$ may be found from this equation by replacing λ by $1/\lambda$, f by g, and interchanging u and v, x_2 and x_3, and a' and b.
The result is

$$(2.10)\quad 2(x_1x_4 - x_2x_3) + 2\lambda x_2^2 + 2(2a'\lambda^2 + \lambda_v)x_2x_4 + (4a'b + 2g\lambda + 2a'_u\lambda + 2a'_v\lambda^2 + 4a'^2\lambda^3 + 6a'\lambda\lambda_v + \lambda_{vv})x_4^2 = 0.$$

3. Properties of the quadrics. A few properties of the quadrics $Q^{(u)}$, $Q^{(v)}$ will now be obtained from a study of their equations. First, however, it is worthy of note that each of these quadrics is defined by a generalization of the definition of the *quadric of Lie*. In fact, the quadric of Lie is the quadric $Q^{(u)}$ in case C_λ is the asymptotic curve C_v, and it is the quadric $Q^{(v)}$ in case C_λ is the asymptotic curve C_u.

Since the equation for the quadric of Lie is

$$(3.1)\qquad x_1x_4 - x_2x_3 + 2a'bx_4^2 = 0,$$

we find that the intersections of the quadric of Lie with the quadric $Q^{(u)}$ are straight lines whose equations are

$$(3.2)\qquad x_3 = x_4 = 0,$$

$$(3.3)\qquad x_3 - \mu_1x_4 = 0, \qquad x_1 - \mu_1x_2 + 2a'bx_4 = 0,$$

and

$$(3.4)\qquad x_3 - \mu_2x_4 = 0, \qquad x_1 - \mu_2x_2 + 2a'bx_4 = 0,$$

where μ_1, μ_2 are the roots of the equation

$$(3.5)\qquad 2\lambda^2\mu^2 + 2\lambda(2b - \lambda_u)\mu + c = 0$$

in which $c = 4b^2 + 2b_u\lambda + 2b_v\lambda^2 + 2f\lambda^2 - 6b\lambda_u + 2\lambda^2_u - \lambda\lambda_{uu}$. The quadric of Lie is, moreover, tangent to $Q^{(u)}$ at each point of the u-tangent $x_3 = x_4 = 0$.

Let π_1, π_2 denote the planes determined by the rulings (3. 2), (3. 3) and (3. 2), (3. 4), respectively. *The planes* π_1, π_2 *are the only two planes each of which is determined by two rulings which are common to both quadrics.*

The harmonic conjugate of the tangent plane to S at P with respect to the planes π_1 and π_2 is the plane whose equation is

$$x_3 + (2b - \lambda_u)x_4/2\lambda = 0. \tag{3. 6}$$

Let N_λ denote the *conjugate net* consisting of the family of curves defined by (2. 3) and the conjugate family defined by

$$du + \lambda du = 0. \tag{3. 7}$$

The *axis of the point P with respect to the net* N_λ joins the point P to the point ζ given by $\zeta = y_{uv} - \alpha y_u - \beta y_v$ wherein $\alpha = (\lambda_v - 2b)/2\lambda^2$, $\beta = -(\lambda_u + 2a'\lambda^3)/2\lambda$. This axis lies in the plane (3. 6) if and only if $a'\lambda^3 + b = 0$. This is the equation of the *directions of Darboux.* Hence, we have the following theorem:

A curve C_λ *is a curve of Darboux if and only if, at a general point P of the curve, the harmonic conjugate of the tangent plane to S at P with respect to the planes* π_1 *and* π_2, *which are determined by the common rulings of the quadric* $Q^{(u)}$ *and the quadric of Lie, contains the axis of P with respect to the conjugate net* N_λ.

In a similar manner the following theorem may be proved:

A curve C_λ *is a curve of Segre if and only if, at a general point P of the curve, the harmonic conjugate of the tangent plane to S at P with respect to the planes* π_1 *and* π_2 *contains the associate axis of P with respect to the conjugate net* N_λ.

The quadric $Q^{(v)}$ may, of course, be used in place of the quadric $Q^{(u)}$ to obtain similar theorems.

Let l denote an arbitrary line in the tangent plane to S at P which generates a congruence Γ as P varies over S. The general homogeneous cöordinates of the points ρ, σ in which l intersects the u and v-tangents to S at P are given by $\rho = y_u - \beta y$, $\sigma = y_v - \alpha y$, where α, β are functions of u, v. Let l' denote the reciprocal of l with respect to the quadric of Lie, and let Γ' denote the congruence generated by l' as P varies over S.

The reciprocal polar lines of l with respect to the quadrics $Q^{(u)}$ and $Q^{(v)}$ are given by the equations

$$x_3 + \beta x_4 = 0, \quad -\lambda^2 x_2 + 2\lambda x_3 + (2b - \lambda_u - \alpha\lambda^2)x_4 = 0, \tag{3. 7}$$

and

$$(3.8)\qquad x_2 + \alpha x_4 = 0,\quad 2\lambda x_2 - x_3 + (2a'\lambda^2 + \lambda_v - \beta)x_4 = 0,$$

respectively. These lines are coplanar with the conjugate of the tangent at P to the curve C_λ if and only if

$$\lambda' = 2b - 2\beta\lambda + 2\alpha\lambda^2 - 2a'\lambda^3.$$

Putting $\lambda = dv/du$, $\lambda' = d^2v/du^2$, we have

$$(3.9)\qquad d^2v/du^2 = 2b - 2\beta dv/du + 2\alpha(dv/du)^2 - 2a'(dv/du)^3.$$

This is the differential equation of the *union-curves of the congruence* Γ'. Hence, we have the theorem:

The curves of a family defined on S by $dv - \lambda du = 0$ are union curves of the congruence Γ' if and only if at a general point P of S the conjugate of the tangent to C_λ at P lies in the plane determined by the reciprocals of l with respect to the parametric osculating quadrics $Q^{(u)}$ and $Q^{(v)}$ of the family at P.

The equations of the tangent planes to the quadrics $Q^{(u)}$, $Q^{(v)}$ at a general point z defined by (2.4) are

$$(3.10)\qquad \lambda^2 x_2 - \lambda x_3 + (\lambda_u - 2b - w\lambda)x_4 = 0,$$

and

$$(3.11)\qquad \lambda x_2 - x_3 + (\lambda_v + 2a'\lambda^2 + w)x_4 = 0,$$

respectively. By making use of these equations and the equations $x_4 = 0$ and $2\lambda(\lambda x_2 - x_3) + (\lambda' - 2b + 2a'\lambda^3)x_4 = 0$ for the tangent plane to S at P and the osculating plane of C_λ at P, the reader may easily establish the property of the quadrics $Q^{(u)}$, $Q^{(v)}$ which is expressed as follows:

THEOREM. *At a general point on the tangent line to C_λ at P the tangent planes to the quadrics $Q^{(u)}$ and $Q^{(v)}$ separate harmonically the osculating plane of C_λ at P and the tangent plane of S at P.*

UNIVERSITY OF KANSAS.

MINIMAL SURFACES SPANNING CLOSED MANIFOLDS AND HAVING PRESCRIBED TOPOLOGICAL POSITION.*

By NORMAN DAVIDS.

The following paper is concerned with minimal surfaces which have both least area and prescribed topological structure, and whose boundary elements belong to prescribed homology-classes on a prescribed closed manifold M.[1]

In [I] and [II] R. Courant has treated variational problems for minimal surfaces with "free boundaries" in a general way, the connectivity being arbitrary and parts of the boundaries free on a manifold, other parts being fixed on prescribed Jordan arcs.[2] However, if the entire boundary of the required minimal surface is left free on the manifold, a new topological condition becomes necessary to exclude trivial degenerate solutions. Such conditions, first mentioned in [II], stipulate linking properties between the boundary elements of the minimal surfaces and certain cycles in the space complementary to the manifold. In [II] and [IV] such linking properties are applied to the discussion of fixed boundary problems; in [III] the following problem is solved: given a closed surface M of genus $p > 0$ in three-space and a closed simple polygon H having no points in common with M, which is linked with all individuals of a non-zero one-dimensional homology-class on M,[3] find a simply-connected minimal surface of least area spanning M whose boundary is linked with H. The solution is based on an interplay of topological and analytical reasoning.

The purpose of this paper is to elaborate the method with some modifications and to extend it to more general problems. We shall allow, in order to formulate comprehensive results, solutions consisting of a number of distinct surfaces. The possibilities arising in these problems far surpass those in the fixed boundary case.

The emphasis in this paper will be on the topological aspects of the problem and the analytical existence proof. It will then follow exactly as in [II] that the solution is a minimal surface and that it satisfies certain transversality conditions on M.

---

* Received April 8, 1941.

[1] See paper [III], where a preliminary report appears and the leading ideas of this paper are summarized.

[2] See [VII], where the problem of free boundaries on planes is treated.

[3] For example, a polygon passing through the hole of a torus.

I. FORMULATION OF THE PROBLEM.

1. **The topology of boundary elements.** Given a closed oriented manifold M[4] of first Betti number $p > 0$ in euclidean n-space E_n having position vector $\mathfrak{x}$: $(x_1, \cdots, x_n)$ the "topological position relative to M" of a surface spanning M is determined by identifying its boundary with prescribed one-dimensional homology-classes on M. The degenerated "point" surface on M then solves the minimum problem for the zero-class; the difficulties in the corresponding existence problem for other homology-classes lie in connecting such prescribed homology conditions with the variational procedure. We shall do this by introducing certain cycles in the complementary space $E_n - M$. If we choose for M an arbitrary one-dimensional homology-basis $h_1, \cdots, h_p$ there exist p oriented $(n-2)$-dimensional cycles $z_1, \cdots, z_p$ in $E_n - M$ forming a homology-basis for $E_n - M$, such that

$$\lambda(h_i, z_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{1}$$

where λ denotes the linking number.[5] The set of cycles may be taken as a finite polyhedral complex Z, of positive distance α from M. Any homology-class on M is equivalent to its set of p linking numbers (1). An orientable surface spanning M will have λ (signed) intersections with the cycle z_i, λ being its linking number with the class of its boundary. These intersections will play an important role in the existence proof.

We now suppose $\mathfrak{x} = \mathfrak{x}(u, v)$ to be a surface defined over the interior of a parameter domain G of the (u, v)-plane bounded by a finite number of continuous curves. Since $\mathfrak{x}$ need not meet M in continuous curves,[6] the above homology or linking conditions referring to boundary curves have to be modified. This is done by replacing the idea of "boundary" by the more general concept of "boundary element," based on the following definition: we shall say that $\mathfrak{x}$ *spans* M if, for all sequences of points (u_n, v_n) in G which tend to a boundary point of G, the distance

$$d[\mathfrak{x}(u_n, v_n), M] \to 0. \tag{2}$$

---

[4] An r-dimensional manifold $(r < n)$ is a continuous, arcwise-connected set locally homeomorphic to an r-sphere. A closed manifold is a compact manifold having no boundary points.

[5] All the topological definitions and theorems used here may be found in [X]. The existence theorem expressed by (1) forms part of the Alexander Duality Theorem.

[6] As pointed out by Courant, the situation may be complicated by the presence of "spikes," i. e., long, thin portions of the surface which wind indefinitely often around the manifold and thus form discontinuities at the boundary.

Hence any sequence of neighborhoods of the boundary B which nest down to B is mapped by $\mathfrak{x}$ into a nested sequence of surface-pieces which converge uniformly to the manifold M. A similar property holds near single boundary curves of B. The class of all sequences of neighborhoods nesting down to a boundary curve of B will be called a *boundary element* of $\mathfrak{x}$ associated with this boundary curve of G.

To each such boundary element of $\mathfrak{x}$ we may associate linking numbers with the cycles $z_1, \cdots, z_p$. For, all curves in G sufficiently near a boundary curve of G and homologous to it have, as images, curves in E_n closer to M than $\alpha/2$, which are therefore homologous in $E_n - \Sigma z_i$. Their common linking number λ_i with z_i will be taken as that of the boundary element. In this manner, by a remark above, we assign to it uniquely a homology-class on M. Our treatment has entirely avoided referring to the intersection of $\mathfrak{x}$ with M.

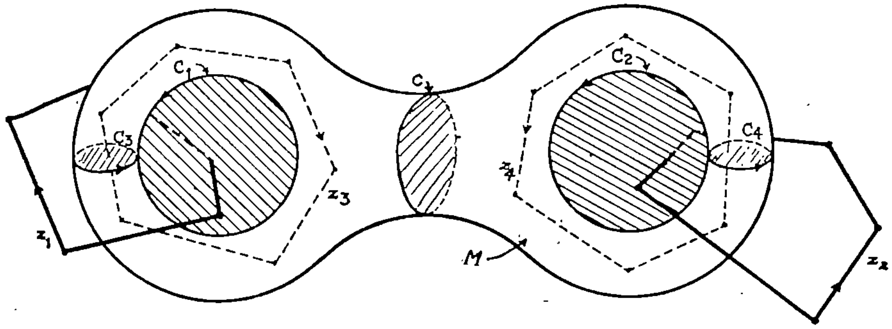


Fig. 1.

In general, if $\mathfrak{x}$ is a surface spanning M having k boundary elements B_j $(j = 1, \cdots, k)$ the kp numbers

$$(3) \qquad \lambda_{ij} = \lambda(B_j, z_i) \qquad i = 1, \cdots, p$$

characterize completely the topological position of $\mathfrak{x}$ relative to M.

The *total boundary* of an orientable surface consists of the totality of its oriented boundary elements. The linking numbers with the z-cycles of the total boundary B are the sum of those for the separate boundary elements.

An example illustrating the theory is that of the two-dimensional manifold M shown in Fig. 1. Here at least four minimal surfaces exist whose boundary elements belong to independent homology-classes.[7]

---

[7] It can be seen intuitively that an additional minimal surface forms at the bottle-like constriction at C (Fig. 1) which, although its area is a relative minimum, cannot be characterized by linking conditions, since its boundary is homologous to zero. Additional conditions of a homotopic nature are required. This illustrates the variety of possibilities which may arise in such free boundary problems.

2. Formulation of the variational problem. A minimal surface[8] $\mathfrak{x} = \mathfrak{x}(u, v)$ defined over a domain G of the (u, v)-plane is obtained by solving, under certain prescribed conditions, the variational problem

$$D(\mathfrak{x}) = \tfrac{1}{2} \int\!\!\int_G (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\, du dv = \min = \delta$$

where $D(\mathfrak{x})$ denotes the classical Dirichlet integral.[9] Its smallest value δ represents the area of the minimal surface.

The possibility of a minimum problem is suggested in a variety of ways. A typical problem would be one requiring a minimal surface spanning M of least area, where the p linking numbers of the total boundary are prescribed, while the topological structure and the number of boundary elements are left free. A more restrictive problem, analogous to the Plateau-Douglas problem,[9a] would specify both the topological structure and the kp integers (3). The solution of the first of these problems gives the key to the second.

A problem requiring interlocking, but with unspecified linking number, always has a solution (see [III]), with an undetermined value of λ. One would then expect that the k-tuply counted solution solves the minimum problem prescribing the linking number $k\lambda$. In fact such multiply-counted surfaces must be permitted in the theory if complete generality of the results is to be attained.

Accordingly, we shall emphasize the concept of a "vector" rather than surface. A vector $\mathfrak{x}$ shall consist of a number of surfaces each defined over its own parameter domain and counted with its proper multiplicity. All the considerations of the previous section apply without change to vectors.

We now introduce the following class of "admissible" vectors: $\mathfrak{x}$ is to consist of a set of surfaces $\mathfrak{x}_r = \mathfrak{x}_r(u, v)$ defined over domains G_r, such that

(a) the components $(x_1, x_2, \cdots)$ of each $\mathfrak{x}_r$ are continuous and have piecewise-continuous derivatives in the interior of G_r,

---

[8] A minimal surface $\mathfrak{x} = \mathfrak{x}(u, v)$ is determined by the conditions

$$\mathfrak{x}_u^2 = \mathfrak{x}_v^2; \quad \mathfrak{x}_u \mathfrak{x}_v = 0$$

and

$$\Delta \mathfrak{x} = 0.$$

[9] The use of the Dirichlet integral, first explicitly introduced by Douglas, forms an important part of both Courant's and Douglas' methods in the fixed boundary case. The above variational problem is similar to Courant's in [II], except that the prescribed conditions are different.

[9a] The classic problem of Plateau for fixed boundaries was first solved by J. Douglas and T. Radó. The extension of the problem to more boundary contours and higher topological structure, proposed and treated by Douglas (see [VI]), has since been called the Plateau-Douglas problem.

(b) the Dirichlet integral of $\mathfrak{x}$ is bounded,

(c) each $\mathfrak{x}_r$ spans M,

(d) the total boundary of $\mathfrak{x}$ satisfies prescribed linking conditions,

(e) each surface of $\mathfrak{x}$ is orientable, and the characteristic $\rho(\mathfrak{x})$ is less than a fixed constant K.[10] Our main theorem then states:

Given a set of p integers $\lambda_1, \cdots, \lambda_p$ not all zero, then, if the class of admissible vectors $\mathfrak{x}$ satisfying the conditions

$$\lambda(B, z_i) = \lambda_i \tag{4}$$

is not empty,[11] *there exists a vector $\mathfrak{x}$ in this class for which the Dirichlet integral attains a minimum value $\delta > 0$. Each surface $\mathfrak{x}_r$ of $\mathfrak{x}$ is a minimal surface and $\rho(\mathfrak{x}) < K$.*

A sufficient condition that $\mathfrak{x}$ consist of a single minimal surface is given by the relation

$$\delta(\lambda_1, \cdots, \lambda_p) \leq \delta(\lambda'_1, \cdots, \lambda'_p) \tag{5}$$

where $\lambda_1, \cdots, \lambda_p$ denote the prescribed linking numbers, $\lambda'_1, \cdots, \lambda'_p$ any other set of p integers, and the notation $\delta(\lambda_1, \cdots, \lambda_p)$ the g. l. b. of the Dirichlet integral considered as a function of the λ's.

3. Preliminary lemmas. We shall require for the existence proof the following lemmas, proved in previous papers:

LEMMA 1. *If $\mathfrak{x} = \mathfrak{x}(u, v)$ is a simply-connected orientable surface spanning M whose boundary element is linked with at least one cycle of Z, then*

$$D(\mathfrak{x}) \geq A \tag{6}$$

where A is a constant independent of $\mathfrak{x}$. This lemma is a corollary to the existence proof of [III].[12]

---

[10] The characteristic of a vector is defined as

$$\rho(\mathfrak{x}) = \Sigma\rho(\mathfrak{x}_r)$$

The condition $\rho < K$ is introduced to prevent the possibility that the characteristics of a minimizing sequence tend to infinity. This case will not be considered here.

[11] If M is piecewise-smooth this condition will always be satisfied.

[12] It is possible to give a direct proof of Lemma 1. The author has obtained the explicit bound

$$A = a^2/8\pi$$

where a is the minimum distance of the z-cycle from M. By using suitable normal domains it is possible to extend the lemma to surfaces of higher connectivity.

The following important convergence theorem on harmonic vectors, due to Courant,[13] will be quoted as

LEMMA 2. *If a sequence of harmonic vectors* $\mathfrak{x}_n$ *with uniformly bounded Dirichlet integrals spans respectively a sequence of manifolds* M_n, *then, if* $\mathfrak{x}_n$ *tends to a limit vector* $\mathfrak{x}$[14] *and* M_n *to a limit manifold* M,[15] $\mathfrak{x}$ *spans* M.

In our applications M_n will be a sequence of curves tending to our given manifold M. Also we shall need the following lemma, due to Courant:[16]

LEMMA 3. *If* $\mathfrak{x}(u,v)$ *is a surface defined over a domain* G *with* $D_G(\mathfrak{x}) < A$, *then, for arbitrarily small* r, *there exists a circle of radius* r' *with* $r \leq r' \leq \sqrt{r}$ *about any point of* G *or the boundary of* G *such that the oscillation of* $\mathfrak{x}$ *on this circle is less than*

$$[4\pi A / |\log r|]^{\frac{1}{2}}.$$

II. SOLUTION OF THE VARIATIONAL PROBLEM.

1. **Normal slit domains.** Before applying the variational procedure, we shall suitably normalize the parameter domains. Any parameter domain must have the same topological structure as the surface it represents, i. e., it must have the same number of boundaries, genus, and orientability character, higher genus being obtained by identifying suitable boundary edges. We shall employ a particular type of normal domain obtained from the theory of conformal mapping, and described as follows:[17] a *normal slit domain* G shall consist of the upper half (u, v)-plane cut by $k-1$ horizontal slits of finite length whose edges, together with the real axis, form the k boundaries of G. In addition, if the genus is p, $2p$ slit pairs are present which extend to $u = +\infty$, the four edges formed by each pair being properly identified. A non-orientable cross-cap is obtained by an additional slit-pair with a different manner of identification. If β is the total number of infinite slit-pairs, the characteristic of the domain is given by

$$\rho = k + \beta - 2. \tag{7}$$

We shall allow also the limiting cases where slits become coalesced.

---

[13] See [I] or [II].

[14] This is understood in the sense that, if $\mathfrak{x}$ is defined over a domain D, $\mathfrak{x}_n$ converges uniformly in every closed subdomain of D.

[15] I. e., the maximum distance of M_n to M tends to zero.

[16] See [V], p. 688.

[17] Such normal domains were first obtained in [IX] pp. 114-122, and in Koebe, *Gott. Nachr.* (1919) pp. 1-46. They are described in detail in [VIII] pp. 855-861.

Our designation of these domains as "normal" domains is justified by the following theorem of conformal mapping: Any domain (or Riemann surface) can be mapped conformally on a suitable normal slit domain.[18] Accordingly, if $\mathfrak{x}(u, v)$ is a surface of finite genus it is always possible to choose new parameters (u, v) which range over a normal slit domain of the same topological structure. We shall henceforth assume that the surfaces of an admissible vector are already represented over such domains. The importance of the conformality of the mapping lies in the fact that the Dirichlet integral is thereby left invariant, and hence the minimum problems remain unaltered.

2. Generalized minimizing sequences. From our assumptions it follows that there exists a minimizing sequence of admissible vectors whose D-integals tend to their greatest lower bound δ. The procedure of the variational problem consists in choosing a convergent minimizing sequence. It must then be established that the limit vector so obtained satisfies all the prescribed conditions, particularly the topological ones. But precisely here lies the difficulty, namely, that of showing that the passage to a limit does not lead out of the class of admissible vectors.

A further difficulty arises due to the irregularity of the surfaces at their boundaries. This will be overcome by the following device, which admits to competition a wider class of vectors: We shall call $\mathfrak{x}_n$ an "admissible sequence" if each member satisfies the admissibility conditions (a) to (e) except that the surfaces of $\mathfrak{x}_n$ need only span manifolds M_n which, as $n \to \infty$, tend to M. We modify this still further by allowing the surfaces of $\mathfrak{x}_n$ to have "holes" whose diameters tend to zero. This may be expressed explicitly by allowing the manifolds M_n to tend, not exactly to M, but to M plus a set of isolated points in space. We define the quantity

$$\delta' = \text{g. l. b. } [\underline{\lim}\, D(\mathfrak{x}_n)] \tag{8}$$

where the g. l. b. is taken for all admissible sequences $\mathfrak{x}_n$. We shall then call $\mathfrak{x}_n$ a *generalized minimizing sequence* (g. m. s.) if $\mathfrak{x}_n$ is an admissible sequence such that

$$D(\mathfrak{x}_n) \to \delta'. \tag{9}$$

Evidently $\delta' \leq \delta$, but we shall prove later that $\delta' = \delta$.

In determining the characteristic of a member of a g. m. s. we shall not count the contributions of its holes contracting to points. This "generalized characteristic" ρ_n is that which the vector $\mathfrak{x}_n$ would have if the holes were not there. We define the quantity

---

[18] The mapping is established by means of a suitable dipole potential on the domain or surface, whose existence is proved by setting up a certain variational problem. The details for genus zero are presented in [XI], and for higher genus in [IX].

$$\underline{\rho} = \text{g. l. b. } [\underline{\lim}\, \rho_n]$$

for all g. m. s. Evidently $\underline{\rho} < K$, and clearly is not greater than the correspondingly lower limit for the ordinary topological characteristic. The two, however, as will be seen later, are actually equal.

Next, $\underline{\rho}$ cannot be $-\infty$. For we may assume that each surface of $\mathfrak{x}_n$ interlocks with a cycle of Z.[19] It follows from Lemma 1 that the total number of simply-connected surfaces of $\mathfrak{x}_n$ cannot exceed $\frac{8\pi A}{\alpha^2}$, where A denotes a uniform upper bound of $D(\mathfrak{x}_n)$. Since the characteristic of each surface of $\mathfrak{x}_n$ must be an integer ≥ -1, we must have $\rho_n \geq -\frac{8\pi A}{\alpha^2}$. We may then choose a g. m. s. $\mathfrak{x}_n$ such that

$$D(\mathfrak{x}_n) \to \mathfrak{d}; \quad \rho_n = \underline{\rho}. \tag{10}$$

Next, denote by G_n the collection of parameter domains belonging to the surfaces of $\mathfrak{x}_n$. It is then possible to choose for each of the domains δ_n-neighborhoods of its boundary B_n, with δ_n so small that the image of the neighborhoods does not meet Z and is nearer to M_n than $1/n$. By deleting these surface-pieces from $\mathfrak{x}_n$ we obtain a new vector $\mathfrak{y}_n$ which, because of (4) and because M_n tends to M, remains an admissible sequence. $\mathfrak{y}_n$ is defined over a set of subdomains G'_n of G_n (formed by deleting the δ_n-neighborhoods from G_n) and its boundary consists of a set of continuous curves C_n. Let $\mathfrak{x}'_n$ be the harmonic vector in G'_n defined by the boundary values

$$\mathfrak{x}'_n(u, v) = \mathfrak{y}_n(u, v) \qquad \text{on } C_n\,{}^{20}$$

Evidently the vectors $\mathfrak{x}'_n$ in G'_n also form an admissible sequence; hence

$$\underline{\lim}\, D_{G'_n}(\mathfrak{x}'_n) \geq \mathfrak{d}.$$

By the classical Dirichlet principle, on the other hand,

$$D_{G'_n}(\mathfrak{x}'_n) \leq D_{G'_n}(\mathfrak{x}_n) \leq D_{G_n}(\mathfrak{x}_n);$$

hence

$$D_{G'_n}(\mathfrak{x}'_n) \to \mathfrak{d},$$

i. e., the $\mathfrak{x}'_n$ also form a g. m. s.

---

[19] Otherwise we could discard the portions which do not interlock and obtain an admissible vector with a smaller D-integral.

[20] This process replaces each surface comprising $\mathfrak{x}$ by a harmonic surface.

This replacement does not alter the generalized characteristics. For we can obtain $\mathfrak{x}'_n$ from $\mathfrak{x}_n$ by means of the deformation represented by the continuous family of vectors

$$\bar{\mathfrak{x}}_n(u,v) = \mathfrak{x}_n(u,v) + t[\mathfrak{x}'_n(u,v) - \mathfrak{x}_n(u,v)] \qquad 0 \leq t \leq 1;$$

hence $\rho'_n \leq \rho_n$,[21] and by (10), $\rho'_n = \rho$.

The next step consists in mapping the domains forming G'_n (which are not slit domains) conformally on normal slit domains, the set of which we again denote by G_n. The surfaces forming $\mathfrak{x}'_n$ map into harmonic surfaces $\mathfrak{x}_n$ over the domains of G_n, and having the same respective D-integrals. By a similarity transformation of the domains we can further arrange that the point $(0,1)$ always corresponds to a Z-intersection.

3. Convergence of the slit-domains. In order to obtain a convergent subsequence of a g. m. s. we must first obtain one all of whose slit-domains converge. We proceed in the following manner: since ρ_n is constant $(=\rho)$ the number of boundary slits in the domains of G_n is ultimately constant, and we can therefore set up a correspondence between all the boundary slits S_n^k of the domains of G_n and those of G_{n+1}.[22] We may then choose a subsequence of $\mathfrak{x}_n$ (and correspondingly a subsequence of G_n) such that the end-points of S_n^k converge, in the (u,v)-plane, to a pair of points lying on the same horizontal line. (A point-pair may coalesce to a single point.) The slit S^k determined by them will be called the limit of S_n^k.[23] We can arrange, in this manner, that all the slits converge for a suitable subsequence. In this way all the slit-domains of G_n define limit slit-domains, the collection of which we denote by G. The domains of G may have degenerate slits (point slits) and, in addition, a set of points which are limits of the holes on which the oscillation of $\mathfrak{x}_n$ tended to zero.

In the course of the convergence we shall prove that the following phenomena cannot occur:

(a) The distance in the (u,v)-plane between two slits tends to zero while each slit remains above a fixed positive length.

(b) A slit or several slits contract to a single point.

---

[21] For this deformation leaves invariant the number of boundary slits, and may, at worst, reduce the number of coördinated slits.

[22] In the case where each domain has characteristic zero, their number may increase indefinitely. A diagonal process then has to be employed to obtain a convergent subsequence.

[23] Except where a slit tends to infinity. If this occurs, the slit is lost in the limit.

For, suppose first that two slits $S_n{}^k$, $S_n{}^{k'}$ both remaining of positive length, were to come arbitrarily near to a point P. Let r_ν be a sequence of positive numbers tending to zero. By Lemma 3 there exists a circle c_ν of radius $r'_\nu < \sqrt{r_\nu}$ such that the oscillation of $\mathfrak{x}_n$ on c_ν is less than $[4\pi A / |\log r_\nu|]^{\frac{1}{2}}$ and hence tends to zero. We take an n_ν so large that the slits $S_{n_\nu}{}^k$, $S_{n_\nu}{}^{k'}$ intersect c_ν; next cut the slit domain of G_{n_ν}, to which these slits belong, along the smaller arc of c_ν joining them. Since the cut domain has one boundary less than the original, its characteristic is

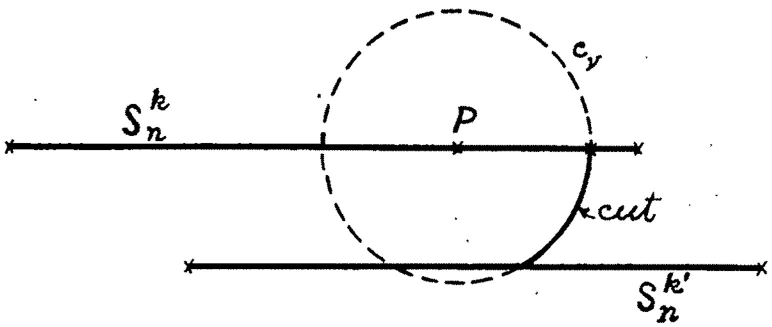


Fig. 2.

$$\rho'_n = \rho - 1.$$

On the other hand, since the boundary of the corresponding cut surface continues to tend to M, the cut surfaces remain a g. m. s. This, however, contradicts (10).

We exclude similarly possibility (b). Suppose one or more slits were to contract to a single point P. Then, exactly as above, we could find a sequence of arcs or circles c_ν contracting to P upon which the oscillation of $\mathfrak{x}_n$ tends to zero, and a suitable subsequence $\mathfrak{x}_{n_\nu} = \mathfrak{x}_\nu$ such that the slits of G_ν contracting to P all lie inside c_ν; this splits the slit domain in question into two domains, the interior and exterior of c_ν. The sum of the ordinary characteristics of the two parts remains equal to that of the original domain, but the total "generalized characteristic" as defined in § 2 is reduced by two, since the curve c_ν defines, for each part, a hole contracting to a point. Thus

$$\rho(\mathfrak{x}_\nu) = \rho - 2$$

which again contradicts (10).

4. Topological position of the limit vector. It becomes possible now to pass immediately to a limit vector. Let G be the limit set of domains as just defined. If any points are present which are the limits of the holes we count these punctures as part of the boundary of G. In any closed subdomain G' interior to G (i. e., subdomains of each slit-domain of G) we can choose a subsequence of $\mathfrak{x}_n$ converging to a limit harmonic vector $\mathfrak{x}$.[24] By a well-known diagonal process applied to a sequence of subdomains tending to G, we may cull a subsequence $\mathfrak{x}_n$ which converges to $\mathfrak{x}$ in the interior of G, uniformly in any closed subdomain G'.

The punctures in the domains of $\mathfrak{x}$ may easily be removed. For, by an elementary theorem of Function theory,[25] since $\mathfrak{x}$ is bounded and regular around these points, we can define values at these points in such a way that the vector becomes harmonic there also. Since the boundary curves of the surfaces of $\mathfrak{x}_n$ tend to M, Lemma 2 gives: *the surfaces of* $\mathfrak{x}$ *span* M. Also

$$D(\mathfrak{x}) \leq \lim D(\mathfrak{x}_n) = \delta';$$

hence

$$D(\mathfrak{x}) \leq \delta'. \tag{11}$$

It remains to be proved that the limit vector $\mathfrak{x}$ satisfies the prescribed linking conditions. This is done by studying the behavior of the points in the parameter domains of G_n which have, as images, intersections with z-cycles. We shall call such points "intersection points" and speak of them as "+" or "—" corresponding to the sign of the image intersection in space. The linking number of $\mathfrak{x}$ with the cycle z_i is then measured by the algebraic sum of these points. If, therefore, this number does not change in the limiting process, it follows that $\lambda(B, z_i) = \lambda_i$, where B denotes the total boundary of $\mathfrak{x}$. A sufficient condition for this is that all the intersection-points of G_n remain at fixed positive distance from all boundary slits. On the other hand a change in the linking number of the limit vector occurs if a set of intersection-points with a non-zero sum tends to boundary slits, these points no longer appearing in the limit domain. By excluding this possibility, we shall complete the proof of the main existence theorem.

Accordingly, suppose that for one cycle z_i it occurs that in the slit-domains G_n a set of intersection-points with a non-zero sum tends to boundary slits. There must then be at least one slit S of G such that the set of intersection-points tending to it have a non-zero sum. Since $\mathfrak{x}$ spans M, given an ϵ arbitrarily small, we can then find a curve c_ϵ encircling S and enclosing a neighborhood S_ϵ of S such that the distance

---

[24] By a well-known theorem of potential theory.

[25] A function analytic around a finite discontinuity may be made analytic there also. This extends immediately to harmonic functions.

(12) $$d[\mathfrak{x}^*, M] < \epsilon; \qquad D(\mathfrak{x}^*) < \epsilon$$

where $\mathfrak{x}^*$ denotes the image, under $\mathfrak{x}$, of the domain S_ϵ. All z_i-intersection points of $\mathfrak{x}$ (which are finite in number) therefore lie in the domain $G'_\epsilon = G - S_\epsilon$. Since $\mathfrak{x}_n$ converges uniformly to $\mathfrak{x}$ in G'_ϵ, we can take n so large that

(13) $$|\mathfrak{x}_n(u, v) - \mathfrak{x}(u, v)| < \epsilon$$

in G'_ϵ, so that the distance of the image of c_ϵ under $\mathfrak{x}_n$ (we denote this image by $C_{n,\epsilon}$) to M is $< 2\epsilon$. The curve $C_{n,\epsilon}$ divides the z_i-intersections of $\mathfrak{x}_n$ into two classes, those in G'_ϵ which converge in the limit to intersections of $\mathfrak{x}$, and those in the complementary domain $S_{n,\epsilon} = G_n - G'_\epsilon$ (the sum of the intersections for both parts being λ_i).

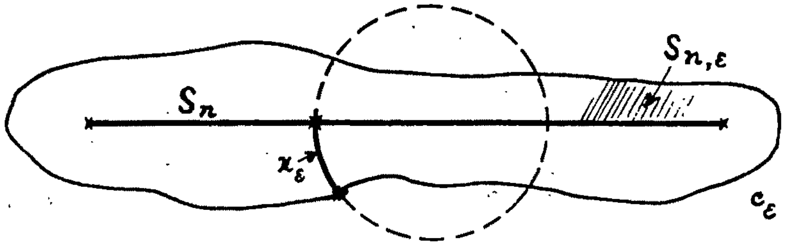


Fig. 3.

For the rest of the reasoning let us suppose, first, that G has at least two boundary slits ($k \geq 2$). Then, if S_n denotes the slit of $\mathfrak{x}_n$ converging to S we can, by Lemma 3, find a circle about any point of S_n with radius r such that $\epsilon \leq r \leq \sqrt{\epsilon}$, and such that the oscillation of $\mathfrak{x}_n$ on this circle tends to zero with ϵ. Let κ_ϵ denote the smallest arc of this circle joining S_n and c_ϵ. We have

(14) $$O(\kappa_\epsilon) = \sigma(\epsilon) \to 0 \text{ as } \epsilon \to 0.$$

We next slit G_n along c_ϵ into two domains ($G_{\epsilon'}$, $S_{n,\epsilon}$) and then slit $S_{n,\epsilon}$ further along the are κ_ϵ into a simply-connected domain. The corresponding split in the vectors $\mathfrak{x}_n$ does not destroy their property of being a g. m. s., because of (12), (13) and (14). The part defined over $S_{n,\epsilon}$ has a non-zero intersection number with z_i, by assumption. The other part, that defined over G'_ϵ, must also have a non-zero intersection number with some z-cycle. Otherwise the intersection numbers of the first part with each z_i would be equal to λ_i, and it would therefore form a g. m. s. by itself, having at least one boundary less than $\mathfrak{x}_n$, again contradicting the fact that $\rho_n = \rho$. Hence the total generalized characteristic is the sum of that for each part, and since these are -1 and ρ

respectively, we obtain a g. m. s. with a characteristic smaller than ρ. This, however, is a contradiction.

We must now dispose of the simply-connected case $k = 1$, where only one slit, the infinite u-axis, is present. If c_ϵ, S_ϵ, G'_ϵ, have the same meaning as before, then two possibilities again arise: either the sum of the intersection-points with at least one z-cycle over G'_ϵ is not zero—in which case the above reasoning applies *verbatim*—or it may happen that the sum of the intersection-points over G'_ϵ vanishes for every z-cycle. In the latter case, however, the portion of $\mathfrak{x}_n$ defined over the complementary domain $S_{n,\epsilon}$ then has the intersection number λ_i with z_i, $i = 1, \cdots, p$, and hence forms a g. m. s. by itself. Therefore

$$D_{S_{n,\epsilon}}(\mathfrak{x}_n) \geq \delta'.$$

On the other hand, since the point $(0, 1)$ because of our normalization remains always an intersection-point with some z-cycle, and since $\mathfrak{x}_n$ is a harmonic vector, we have, by the remark after Lemma 1,

$$D_{G'_\epsilon}(\mathfrak{x}_n) \geq \alpha^2/8\pi;$$

hence

$$D_{S_{n,\epsilon}}(\mathfrak{x}_n) = D_{G_n} - D_{G'_\epsilon} < D_{G_n} - \alpha^2/8\pi < \delta',$$

which is a contradiction. With this our theorem is proved, namely: *The limit vector* $\mathfrak{x}$ *satisfies the prescribed topological conditions,* i. e.,

$$\lambda(B, z_i) = \lambda_i, \qquad i = 1, \cdots, p.$$

Further, since no boundary slits or surface-pieces are lost by the convergence of the slit domains, $\rho(\mathfrak{x}) \leq \rho$, hence it follows by definition of ρ that $\rho(\mathfrak{x}) = \rho$. Since $\mathfrak{x}$ spans M, it is a vector in the original vector class defining δ ([I], § 2), hence $D(\mathfrak{x}) \geq \delta$. (11) then gives

$$\delta \leq D(\mathfrak{x}) \leq \delta',$$

hence, on account of $\delta \geq \delta'$,

$$D(\mathfrak{x}) = \delta = \delta'.$$

Thus $\mathfrak{x}$ solves the minimum problem; and with this the existence proof is completed. At the same time we have proved $\delta = \delta'$, which shows that the extension of the vector-class made in defining δ' did not improve the value of the lower bound δ.

The inequality (5) gives a sufficient condition that $\mathfrak{x}$ consist of a single surface. For, suppose it were possible to divide $\mathfrak{x}$ into at least two separated parts $\mathfrak{x}'$, $\mathfrak{x}''$. Then to each part would be associated its own set of linking numbers $\lambda'_1, \cdots, \lambda'_p$, $\lambda''_1, \cdots, \lambda''_p$ respectively in such a way that $\lambda_i = \lambda'_i + \lambda''_i$

$i = 1, \cdots, p$, and such that both sets differ in at least one number from the original set $\{\lambda_i\}$. (Otherwise one of the surface-parts would not interlock with any cycle z_i, and would therefore be a constant.) Thus

$$D(\mathfrak{x}') = D(\mathfrak{x}) - D(\mathfrak{x}'') < D(\mathfrak{x})$$

and we would have

$$\delta(\lambda'_1, \cdots, \lambda'_p) < \delta(\lambda_1, \cdots, \lambda_p)$$

for this set of λ''s, which is a contradiction. This proves the sufficiency of (5).

That each surface of $\mathfrak{x}$ is a minimal surface follows from the fact that each of these must have smallest Dirichlet integral with respect to all surfaces spanning M having its own topological position. We may exploit this minimum property and introduce certain cuts in the parameter domains, performing variations on the coordinations of the values along the edges of the cuts exactly as done in the Plateau-Douglas problem for fixed boundaries.[26]

5. Concluding remarks and generalizations.[27] The solution just found answers the problem in which a minimal surface spanning M is sought as before, but having also a prescribed topological structure. Its existence conditions are expressible in the form of inequalities similar to those of the Plateau-Douglas problem.

When non-orientable surfaces are admitted to the variational problem [28] intersection numbers are not defined, hence an additional reasoning is required to show the non-constancy of the solution.

In the special space E_3 we have the following further results: a suitable homology-basis exists on M and a set of minimal surfaces $H_1, \cdots, H_p$ such that

$$\lambda(B_i, z_j) = \begin{matrix} 1 \\ 0 \end{matrix} \left\{ \begin{matrix} i = j & i = 1, \cdots, p \\ i \neq j & j = 1, \cdots, p \end{matrix} \right.$$

where $B_i = B(H_i)$ and $z_1, \cdots, z_p$ form the complementary basis in $E_3 - M$.

NEW YORK UNIVERSITY.

REFERENCES.

I. R. Courant, "The existence of a minimal surface of least Area bounded by prescribed Jordan arcs and prescribed surfaces," *Proceedings of the National Academy of Sciences*, vol. 24 (1938), pp. 97-101.

---

[26] See [V], pp. 717-720.

[27] The results mentioned in this section will be treated in detail in a forthcoming paper by the author.

[28] Such minimal surfaces may arise, for example, when M is a knotted torus.

II. R. Courant, "The existence of minimal surfaces under prescribed boundary conditions," *Acta Mathematica*, vol. 72 (1940), pp. 51-98.

III. Courant & Davids, "Minimal surfaces spanning closed manifolds," *Proceedings of the National Academy of Sciences*, vol. 26 (1940), pp. 194-199.

IV. R. Courant, "On a generalized form of the problem of Plateau," to appear soon.

V. R. Courant, "Plateau's problem and Dirichlet's principle," *Annals of Mathematics*, vol. 38 (1937), pp. 679-724.

VI. J. Douglas, "Minimal surfaces of higher topological structure," *Annals of Mathematics*, vol. 40 (1939), pp. 205-298.

VII. I. Ritter, "Solution of Schwarz' problem concerning minimal surfaces," not yet published.

VIII. M. Shiffman, "The problem of Plateau for minimal surfaces of arbitrary topological structure," *American Journal of Mathematics*, vol. 60 (1938), pp. 853-882.

IX. R. Courant, "Über konforme Abbildung von Bereichen, welche nicht durch alle Ruckehrschnitte zerstückelt werden, auf schlichte Normalbereiche," *Mathematische Zeitschrift*, vol. 3 (1919), pp. 114-122.

X. Alexandroff-Hopf, *Topologie* I, pp. 413-425.

XI. Hurwitz-Courant, *Funktionentheorie.*

For a comprehensive bibliography of the literature on the Plateau and related problems, see [II].

AN EXTENSION THEORY FOR RINGS.*

By C. J. Everett, Jr.**

1. Introduction.

As in the theory of groups, many of the more intricate questions of ring structure appear as extension problems. The structure of nilpotent algebras, and that of algebras with given radical and semi-simple residue class algebra modulo this radical serve as examples. It is the purpose of this paper to provide a ring-theoretic analogue of the O. Schreier "Erweiterungstheorie" for groups,[1] and to construct by extensions rings whose chains of left and right annihilators for the successive powers of the ring have their first repetitions at arbitrary positions.

2. Extensions and factor systems.

If $\mathfrak{R} = \{0, A, B, \cdots\}$ is a ring, the set $\Omega(\mathfrak{R})$ of all operators Θ on the additive group of $\mathfrak{R}$ $(\Theta(A+B) = \Theta A + \Theta B,\ A, B \in \mathfrak{R})$, with equality $\Theta = \Gamma$, addition $\Theta + \Gamma$, multiplication $\Theta\Gamma$ defined by $\Theta A = \Gamma A$, $(\Theta + \Gamma)A = \Theta A + \Gamma A$, $(\Theta\Gamma)A = \Theta(\Gamma A)$, (all $A \in \mathfrak{R}$) respectively, is a ring (the endomorphism ring[2] of the additive group of $\mathfrak{R}$). An element A acting as left and as right multiplier on $\mathfrak{R}$ defines operators $\Lambda(A)$, $\mathrm{P}(A)$ in $\Omega(\mathfrak{R})$ thus: $A\mathfrak{R} = \Lambda(A)\mathfrak{R}$. $\mathfrak{R}A = \mathrm{P}(A)\mathfrak{R}$ elementwise.[3] Denoting by $\Lambda(\mathfrak{R})$ the set of all $\Lambda(A)$, $A \in \mathfrak{R}$, and by $\mathrm{P}(\mathfrak{R})$ all $\mathrm{P}(A)$, $A \in \mathfrak{R}$, it is well known that, under the correspondences $A \rightarrow \Lambda(A)$, $A \rightarrow \mathrm{P}(A)$, one has $\mathfrak{R} \sim \Lambda(\mathfrak{R})$, $\mathfrak{R} \backsim \mathrm{P}(\mathfrak{R})$.[4] The subrings $\Lambda(\mathfrak{R})$, $\mathrm{P}(\mathfrak{R})$ of $\Omega(\mathfrak{R})$ are called the *left* and *right regular representation rings of* $\mathfrak{R}$.[5] The correspondences are one-one if and only if the left and right annihilators of $\mathfrak{R}$ are zero.

If $\mathfrak{R} = \{0, A, B, \cdots\}$, $\mathfrak{F} = \{o, \sigma, \tau, \cdots\}$ are rings, and a ring $\mathfrak{E}$ contains the ideal $\mathfrak{R}$, with (i) $\mathfrak{E}/\mathfrak{R} \cong \mathfrak{F}$, then $\mathfrak{E}$ is called an $\mathfrak{R}$—$\mathfrak{F}$-extension. If $\mathfrak{E}$ is such an extension, write $\mathfrak{E} = \Sigma(S_\sigma + \mathfrak{R})$ (S_σ a complete representation

---

* Abstract 47-1-4 *Bulletin of the American Mathematical Society*; received March 6, 1941.

** Sterling Fellow, Yale University, 1940-41.

[1] H. Zassenhaus, *Lehrbuch der Gruppentheorie* I (1937), pp. 89-98. O. Schreier, "Über die Erweiterung von Gruppen," Teil I. *Monats. f. Math. u. Phys.*, vol. 34, pp. 165-180. Teil II. *Hamb. Abh.*, vol. 4, pp. 321-346.

[2] van der Waerden, *Moderne Algebra* I, 2nd ed. (1937), p. 146.

[3] All operators are arbitrarily written on the left.

[4] $R \sim S$ denotes ring homomorphism. $R \backsim S$ means addition is preserved, and $r \rightarrow s$, $r' \rightarrow s'$ implies $rr' \rightarrow s's$.

[5] C. J. Everett, "Annihilator ideals and representation iteration for abstract rings," *Duke Mathematical Journal*, vol. 5 (1939), pp. 623-627.

system[6] for $\mathfrak{E} \bmod \mathfrak{N}$) and in (i) $S_\sigma + \mathfrak{N} \leftrightarrow \sigma$. Since $\mathfrak{N}$ is an ideal in $\mathfrak{E}$, the S_σ acting as left and right multipliers on $\mathfrak{N}$ define operators Λ_σ, P_σ in $\Omega(\mathfrak{N})$: $S_\sigma\mathfrak{N} = \Lambda_\sigma\mathfrak{N}$, $\mathfrak{N}S_\sigma = \mathrm{P}_\sigma\mathfrak{N}$ elementwise; and by virtue of (i), $S_\sigma + S_\tau = S_{\sigma+\tau} + A_{\sigma,\tau}$, $S_\sigma S_\tau = S_{\sigma\tau} + M_{\sigma,\tau}$, $A_{\sigma,\tau}$, $M_{\sigma,\tau} \in \mathfrak{N}$. The extension $\mathfrak{E}$ with representation system S_σ thus defines a set $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$ satisfying the following conditions:

1. $A_{\sigma,\tau} = A_{\tau,\sigma}$.
2. $A_{\sigma,\tau} + A_{\sigma+\tau,\rho} = A_{\tau,\rho} + A_{\sigma,\tau+\rho}$.
3. $A_{\sigma,o} = 0$.
4. $\Lambda_\sigma \mathrm{P}_\tau = \mathrm{P}_\tau \Lambda_\sigma$; $\Lambda_\sigma \mathrm{P}(A) = \mathrm{P}(A)\Lambda_\sigma$; $\mathrm{P}_\sigma \Lambda(A) = \Lambda(A)\mathrm{P}_\sigma$.
5. $\Lambda(A)\Lambda_\sigma = \Lambda(\mathrm{P}_\sigma A)$.
6. $\Lambda_\sigma + \Lambda_\tau = \Lambda_{\sigma+\tau} + \Lambda(A_{\sigma,\tau})$; $\Lambda_\sigma\Lambda_\tau = \Lambda_{\sigma\tau} + \Lambda(M_{\sigma,\tau})$.
7. $\mathrm{P}_\sigma + \mathrm{P}_\tau = \mathrm{P}_{\sigma+\tau} + \mathrm{P}(A_{\sigma,\tau})$; $\mathrm{P}_\sigma \mathrm{P}_\tau = \mathrm{P}_{\tau\sigma} + \mathrm{P}(M_{\tau,\sigma})$.
8. $\Lambda_o = 0 = \mathrm{P}_o$; $M_{o,\sigma} = 0 = M_{\sigma,o}$.
9. $\Lambda_\sigma M_{\tau,\rho} + M_{\sigma,\tau\rho} = \mathrm{P}_\rho M_{\sigma,\tau} + M_{\sigma\tau,\rho}$.
10. $M_{\sigma,\tau+\rho} + \Lambda_\sigma A_{\tau,\rho} = M_{\sigma,\tau} + M_{\sigma,\rho} + A_{\sigma\tau,\sigma\rho}$.
    $M_{\sigma+\tau,\rho} + \mathrm{P}_\rho A_{\sigma,\tau} = M_{\sigma,\rho} + M_{\tau,\rho} + A_{\sigma\rho,\tau\rho}$.

Abstracting from this, with $\mathfrak{N}$, $\mathfrak{F}$ arbitrary rings, any set $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$, Λ_σ, P_σ in $\Omega(\mathfrak{N})$, $A_{\sigma,\tau}$, $M_{\sigma,\tau}$ in $\mathfrak{N}$, satisfying 1-10 is called an $\mathfrak{N}-\mathfrak{F}$-*factor set*. Thus the extension $\mathfrak{E}$ with S_σ defines a factor set $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$. If $T_\sigma = S_\sigma + A_\sigma$, $A_\sigma \in \mathfrak{N}$, is a second representation system for $\mathfrak{E} \bmod \mathfrak{N}$ then T_σ defines a factor set $\{\Lambda'_\sigma, \mathrm{P}'_\sigma, A'_{\sigma,\tau}, M'_{\sigma,\tau}\}$ related to the first by

E1. $\Lambda'_\sigma = \Lambda_\sigma + \Lambda(A_\sigma)$.
E2. $\mathrm{P}'_\sigma = \mathrm{P}_\sigma + \mathrm{P}(A_\sigma)$.
E3. $A_0 = 0$.
E4. $A'_{\sigma,\tau} = A_\sigma + A_\tau + A_{\sigma,\tau} - A_{\sigma+\tau}$.
E5. $M'_{\sigma,\tau} = M_{\sigma,\tau} + \Lambda_\sigma A_\tau + \mathrm{P}_\tau A_\sigma + A_\sigma A_\tau - A_{\sigma\tau}$.

Again abstracting, any two $\mathfrak{N}-\mathfrak{F}$-factor sets related by elements A_σ of $\mathfrak{N}$ through E1-E5 are called *equivalent*; notation: $\{\Lambda_\sigma, \cdots\} \sim \{\Lambda'_\sigma, \cdots\}$. This is an equivalence relation, and by a well known principle[7] one may speak of equivalence classes of $\mathfrak{N}-\mathfrak{F}$-factor sets.

The particular representation system of an extension may thus be ignored; one says that every $\mathfrak{N}-\mathfrak{F}$-extension defines an equivalence class C of $\mathfrak{N}-\mathfrak{F}$-factor sets. Hence the correspondence:

---

[6] Representation system will always signify *normality*: the representative S_0 of the residue class $\mathfrak{N}$ being taken as zero.

[7] van der Waerden I, p. 12.

(ii) $$\mathfrak{E} \to C = (\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\})$$

If $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$ is any $\mathfrak{N} - \mathfrak{F}$-factor set, the set $\mathfrak{P}$ of all formal pairs (σ, A), $\sigma \in \mathfrak{F}$, $A \in \mathfrak{N}$, subject to the definitions

P1. $(\sigma, A) = (\tau, B)$ means $\sigma = \tau$, $A = B$.
P2. $(\sigma, A) + (\tau, B) = (\sigma + \tau, A + B + A_{\sigma,\tau})$.
P3. $(\sigma, A) \cdot (\tau, B) = (\sigma\tau, \Lambda_\sigma B + \mathrm{P}_\tau A + AB + M_{\sigma,\tau})$.

is an $\mathfrak{N} - \mathfrak{F}$-extension containing $(o, \mathfrak{N}) \cong \mathfrak{N}$, and, with representation system $(\sigma, 0)$, defines the factor set $\{\bar{\Lambda}_\sigma, \bar{\mathrm{P}}_\sigma, (o, A_{\sigma,\tau}), (o, M_{\sigma,\tau})\}$, $\bar{\Lambda}_\sigma(o, A) = (o, \Lambda_\sigma A)$, $\bar{\mathrm{P}}_\sigma(o, A) = (o, \mathrm{P}_\sigma A)$. Hence *every class of $\mathfrak{N} - \mathfrak{F}$-factor sets is defined in (ii) by some extension.*

If $\mathfrak{E}$ with S_σ defines $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$ and $\mathfrak{P}$ is the ring defined by P1-3, then $\mathfrak{E} \cong \mathfrak{P}$ under $S_\sigma + A \leftrightarrow (\sigma, A)$. Thus, *if in (ii) $\mathfrak{E}_1 \to C$ and $\mathfrak{E}_2 \to C$, then $\mathfrak{E}_1 \cong \mathfrak{E}_2$.*

3. Direct sums and splitting rings. For any rings $\mathfrak{N}$, $\mathfrak{F}$, there always exists the direct sum[8] extension $\mathfrak{F} + \mathfrak{N}$, defining the $\mathfrak{N} - \mathfrak{F}$-factor set $\{0, 0, 0, 0\}$. An $\mathfrak{N} - \mathfrak{F}$-extension is said to *split* over $\mathfrak{N}$ in case there exists a representation system $\{T_\sigma\}$ which forms a ring: $T_\sigma + T_\tau = T_{\sigma+\tau}$, $T_\sigma T_\tau = T_{\sigma\tau}$. The $\mathfrak{N} - \mathfrak{F}$-factor set class defined by such an extension must contain a set of form $\{\Lambda_\sigma, \mathrm{P}_\sigma, 0, 0\}$: *If an extension $\mathfrak{E}$ with representation system S_σ defines $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$ then $\mathfrak{E}$ splits over $\mathfrak{N}$ if and only if there exists a set of A_σ in $\mathfrak{N}$ satisfying* E3, *and* E4, E5 *with left member zero.*

Let $\mathfrak{E}$ be an $\mathfrak{N} - \mathfrak{F}$-extension with $\mathfrak{E}/\mathfrak{N} \cong \mathfrak{F}$, $\mathfrak{E} = \Sigma(S_\sigma + \mathfrak{N})$ in the previous sense. A ring $\mathfrak{C}$ is said to be a *splitting ring for* $\mathfrak{E}$ mod $\mathfrak{N}$ *relative to* S_σ, in case $\mathfrak{E}$ may be embedded (in the sense of isomorphism) in $\mathfrak{C}$, $\mathfrak{C}$ contains an ideal $\bar{\mathfrak{E}}$ which contains $\mathfrak{N}$, S_σ is a representation system for $\mathfrak{C}$ mod $\bar{\mathfrak{E}}$, and $\mathfrak{C}$ splits over $\bar{\mathfrak{E}}$.

Such a splitting ring always exists. For let $\mathfrak{C} = \mathfrak{F} + \mathfrak{E}$, direct sum, elements $(\sigma, S_\tau + A)$. Then let $\mathfrak{E}'$ be the subring of elements $(\sigma, S_\sigma + A)$ isomorphic under $(\sigma, S_\sigma + A) \leftrightarrow S_\sigma + A$ with $\mathfrak{E}$. Under this correspondence $\mathfrak{N}' = (o, \mathfrak{N}) \cong \mathfrak{N}$, and $(\sigma, S_\sigma) \leftrightarrow S_\sigma$. Clearly $\mathfrak{C} > \mathfrak{E}' > \mathfrak{N}'$. Let $\bar{\mathfrak{E}}$ be the subring of elements $(o, S_\tau + A)$. Then $\mathfrak{C} > \bar{\mathfrak{E}}$ ideal $> \mathfrak{N}'$. $\mathfrak{C}$ splits over $\bar{\mathfrak{E}}$. Finally (σ, S_σ) is a complete representation system for $\mathfrak{C}$ mod $\bar{\mathfrak{E}}$, as required.

Exactly analogous remarks hold for groups.[9] If $\mathfrak{G} > \mathfrak{N}$ as invariant subgroup, $\mathfrak{G}/\mathfrak{N} \cong \mathfrak{F}$, $\mathfrak{G} = \Sigma(\mathfrak{N}S_\sigma)$, then $\mathfrak{E} = \mathfrak{F} \times \mathfrak{G}$ (direct product), elements (σ, AS_τ) has the properties: $\mathfrak{E} > \mathfrak{G}'$ where $\mathfrak{G}'$ is the subgroup of all elements (σ, AS_σ), $\mathfrak{G}' \cong \mathfrak{G}$ under $(\sigma, AS_\sigma) \leftrightarrow AS_\sigma$. $\mathfrak{G}' > \mathfrak{N}' =$ set $(\epsilon, \mathfrak{N})$. $\mathfrak{E} > \bar{\mathfrak{G}} = (\epsilon, \mathfrak{G})$ as invariant subgroup and $\bar{\mathfrak{G}} > \mathfrak{N}'$; $\mathfrak{E}$ splits over $(\epsilon, \mathfrak{G})$. Finally, (σ, S_σ) is a complete representation system for $\mathfrak{E}$ mod $\mathfrak{G}$.

---

[8] If R, S are rings, $R + S$ designates ring-direct sum, with R, S ideals in the sum.

[9] Cf. H. Zassenhaus, *loc. cit.*[1], p. 98, Iyanaga, "Zum Beweis des Hauptidealsatz," *Hamb. Abh.*, vol. 10 (1934), pp. 349-357. This proof requires $\mathfrak{N}$ abelian, and constructs $\mathfrak{N}'$ abelian, essential for the class field application. The present construction satisfies

4. A restricted class group. Let $\mathfrak{N}$ be a null ring ($AB = 0$, all A, B in $\mathfrak{N}$); then $\Lambda(\mathfrak{N}) = (0) = \mathrm{P}(\mathfrak{N})$, $\mathfrak{F}$ arbitrary. Consider the set $\{C\}$ of all classes of $\mathfrak{N} - \mathfrak{F}$-factor sets, $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$. Here $\mathfrak{F} \sim \{\Lambda_\sigma\}$, $\mathfrak{F} \backsim \{\mathrm{P}_\sigma\}$ under $\sigma \to \Lambda_\sigma$, $\sigma \to \mathrm{P}_\sigma$, and $\{\Lambda_\sigma\}$, $\{\mathrm{P}_\sigma\}$ are elementwise commutative rings. Note that $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\} \sim \{\Lambda'_\sigma, \mathrm{P}'_\sigma, A'_{\sigma,\tau}, M'_{\sigma,\tau}\}$ implies $\Lambda'_\sigma = \Lambda_\sigma$, $\mathrm{P}'_\sigma = \mathrm{P}_\sigma$ (E1-2) so that factor sets in the same class have the same operator systems. If, conversely, $\{\Lambda_\sigma\}$, $\{\mathrm{P}_\sigma\}$ are any elementwise commutative sets in $\Omega(\mathfrak{N})$ such that $\mathfrak{F} \sim \{\Lambda_\sigma\}$, $\mathfrak{F} \backsim \{\mathrm{P}_\sigma\}$ under $\sigma \to \Lambda_\sigma$, $\sigma \to \mathrm{P}_\sigma$, then $\{\Lambda_\sigma, \mathrm{P}_\sigma, 0, 0\}$ is an $\mathfrak{N} - \mathfrak{F}$-factor set. Now denote by $\{C'\}$ the set of classes C' whose factor sets $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\}$ contain a fixed pair $\{\Lambda_\sigma\}$, $\{\mathrm{P}_\sigma\}$ of such operator rings. The set $\{C'\}$ is not void since the C' containing $\{\Lambda_\sigma, \mathrm{P}_\sigma, 0, 0\}$ belongs to it. Now define $\{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau}, M_{\sigma,\tau}\} + \{\Lambda_\sigma, \mathrm{P}_\sigma, A'_{\sigma,\tau}, M'_{\sigma,\tau}\} = \{\Lambda_\sigma, \mathrm{P}_\sigma, A_{\sigma,\tau} + A'_{\sigma,\tau}, M_{\sigma,\tau} + M'_{\sigma,\tau}\}$ as the *sum* of two such factor sets. This is again an $\mathfrak{N} - \mathfrak{F}$-factor set, and it is trivial that the addition is well defined relative to the equivalence classes C'. Hence $C'_1 + C'_2$ is well defined as the class C'_3 containing any of the sums of factor sets in C'_1, C'_2.

The set $\{C'\}$ is an abelian group, the unit class $0'$ being that C' which contains $\{\Lambda_\sigma, \mathrm{P}_\sigma, 0, 0\}$. Thus, *an abelian group is formed by the classes of* $\mathfrak{N}$ *(null-ring)* $-\mathfrak{F}$*-factor sets containing any fixed elementwise commutative operator rings* $\{\Lambda_\sigma\}$, $\{\mathrm{P}_\sigma\}$ *with* $\mathfrak{F} \sim \{\Lambda_\sigma\}$, $\mathfrak{F} \backsim \{\mathrm{P}_\sigma\}$ *under* $\sigma \to \Lambda_\sigma$, $\sigma \to \mathrm{P}_\sigma$.[10]

The existence of a single $\mathfrak{N} - \mathfrak{F}$-extension which does not split over $\mathfrak{N}$ will insure the existence of groups $\{C'\}$ having more than one element. It may be remarked here that if $\mathfrak{N}$ is a null ring ($\mathfrak{N}^2 = 0$), and $\mathfrak{F}$ has a right unit ϵ, then any $\mathfrak{N} - \mathfrak{F}$-extension $\mathfrak{E}$ having a representation system S_σ, mod $\mathfrak{N}$, for which $\mathfrak{N}S_\sigma = 0$, $\sigma \in \mathfrak{F}$, must split over $\mathfrak{N}$. (Use the new system $T_\sigma = S_\sigma + A_\sigma$, with $A_\sigma = M_{\sigma,\epsilon}$. The T_σ may be verified to form a ring directly, or use may be made of (10) in E3-5 of § 2.)

5. Closed rings. Holomorph. The parallelism between group and ring extension may be developed further. Let $\mathfrak{N}$ be a ring, $\Lambda = \Lambda(\mathfrak{N})$, $\mathrm{P} = \mathrm{P}(\mathfrak{N})$

---

the Zassenhaus definition for splitting group and is general. See the footnote, Zassenhaus, p. 98.

[10] This is the analogue of the group case with $\mathfrak{N}$ an abelian invariant subgroup. Cf. Zassenhaus, *loc. cit.*[1], p. 93.

the regular representation rings of $\mathfrak{N}$ in $\Omega(\mathfrak{N})$, and denote by P' the $\Omega(\mathfrak{N})$ commutator of P, i. e., the set of all Γ in $\Omega(\mathfrak{N})$ such that $\Gamma\mathrm{P}(A) = \mathrm{P}(A)\Gamma$, all $A \in \mathfrak{N}$. Due to associativity in $\mathfrak{N}$, $\Lambda \leqq \mathrm{P}'$ and Λ *is a left ideal in* P', since (ii) $\Gamma\Lambda(A) = \Lambda(\Gamma A)$, $\Gamma \in \mathrm{P}'$, $A \in \mathfrak{N}$. Let $\mathfrak{L}$ be the left annihilator of $\mathfrak{N}$: $\mathfrak{L} = \{A \in \mathfrak{N};\ A\mathfrak{N} = 0\}$. In what follows, the concepts "group centrum" and "left annihilator," "automorphism group of a group" and "ring P' of all P-operator endomorphisms of $\mathfrak{N}$," and "inner automorphism of a group" and "endomorphism of $\mathfrak{N}$ defined by a left multiplier" play strikingly similar rôles.

A ring $\mathfrak{N}$ is said to be *closed* in case $\mathfrak{L} = (0)$, and $\Lambda = \mathrm{P}'$.

THEOREM. *If $\mathfrak{E}$ contains the closed ring $\mathfrak{N}$ as ideal, then $\mathfrak{E} = \mathfrak{N} + \mathfrak{N}_0$ where $\mathfrak{N}_0$ is the left annihilator of $\mathfrak{N}$ in $\mathfrak{E}$.*[11]

$\mathfrak{N}_0 = \{R \in \mathfrak{E};\ R\mathfrak{N} = 0\}$ is obviously a two-sided ideal. $\mathfrak{N} \cap \mathfrak{N}_0 = (0)$ since for $\mathfrak{N}$, $\mathfrak{L} = (0)$. Let $R \in \mathfrak{E}$. Since $R(N_1 + N_2) = RN_1 + RN_2$, $(RN_1)N_2 = R(N_1N_2)$, $R \in \mathfrak{E}$, R defines as left multiplier on $\mathfrak{N}$ an operator in P', where $\mathrm{P} = \mathrm{P}(\mathfrak{N})$. Since $\Lambda = \mathrm{P}'$ for $\mathfrak{N}$, there is an $A \in \mathfrak{N}$ such that $RN = AN$, all N in $\mathfrak{N}$. Hence $R = A + N'$, $N' \in \mathfrak{N}_0$.

The theorem loses its strange appearance[12] if one notes the

THEOREM. *A ring $\mathfrak{N}$ is closed if and only if it has a unit. For such a ring $\mathfrak{N} \cong \mathrm{P}'$.*[13]

One may prove first a lemma of some interest in itself:

LEMMA. *A ring $\mathfrak{N}$ has a left unit E_l if and only if $\Lambda = \mathrm{P}'$.*[14]

If E_l is a left unit for $\mathfrak{N}$, $\Lambda(E_l) = \epsilon$, unit operator of $\Omega(\mathfrak{N})$, and by (ii) above, $\Gamma\Lambda(E_l) = \Gamma = \Lambda(\Gamma E_l)$, $\Gamma \in \mathrm{P}'$. Thus $\mathrm{P}' = \Lambda$. Conversely, if $\Lambda = \mathrm{P}'$, then ϵ is in P', hence in Λ and for some A, $\Lambda(A) = \epsilon$. Then A is a left unit for $\mathfrak{N}$.

The theorem now follows. For if $\mathfrak{N}$ is closed, $\mathfrak{N} \cong \Lambda = \mathrm{P}'$, and some $A \to \epsilon$ in $\mathrm{P}' = \Lambda$. Since $\mathfrak{N} \cong \Lambda$, A is a unit in $\mathfrak{N}$. (Or, from the lemma, closure of $\mathfrak{N}$ implies a left unit E; all left units are of the form $E + L$, $L \in \mathfrak{L}$.

---

[11] Note that the proof is the exact analogue of that for groups: A. Speiser, *Theorie der Gruppen von endlicher Ordnung*, 2nd ed. (1927), p. 127, Satz 110.

[12] It is proved for algebras, using a Pierce decomposition, by A. A. Albert, *Structure of Algebras, A. M. S. Colloquium Publications* (1939), p. 30, Theorem 23.

[13] This is the essential point in the van der Waerden II treatment of ring theory, in which a (semi-simple) ring with unit is regarded as isomorphic to the endomorphism ring of a group with operators P.

[14] For linear algebras, cf. C. C. MacDuffee, "Modules and Ideals in a Frobenius Algebra," *Monats. f. Math. u. Phys.*, vol. 48 (1939), Theorem 3. 1.

Since $\mathfrak{L} \neq (0)$, E is a unique left unit. But a unique left unit is a unit.) Conversely, if $\mathfrak{N}$ has a unit E, then from the lemma, $\Lambda = \mathrm{P}'$, and obviously $\mathfrak{L} = (0)$, since $\mathfrak{L}E = \mathfrak{L} = (0)$.

It is natural to question the existence of over-rings $\mathfrak{E}$ containing a given ring $\mathfrak{N}$ as ideal, and such that every operator in P' $(\mathrm{P} = \mathrm{P}(\mathfrak{N}))$ is defined by a left multiplier in $\mathfrak{E}$. This is not true in general. For such an over-ring one has $\mathfrak{N}(\mathfrak{E}N) = (\mathfrak{N}\mathfrak{E})N$, all $N \epsilon \mathfrak{N}$, elementwise, whence $\Lambda(\mathfrak{N})\mathrm{P}' \leqq \Lambda(\mathfrak{N})$, and Λ is a two-sided ideal in P'. However for the algebra $e_1{}^2 = e_1$, $e_1e_2 = e_2{}^2 = 0$, $e_2e_1 = e_2$, P consists of the scalar 2×2 matrices, P' is the total 2×2 matrix algebra, $\Lambda = \left(\begin{smallmatrix} a & 0 \\ b & 0 \end{smallmatrix}\right)$, and $\Lambda\mathrm{P}'$ is not in Λ.

However, one may state the following

THEOREM. *Let $\mathfrak{N}$ be a ring with $\mathfrak{L} = (0)$ and let Λ be a two-sided ideal in P'. In particular, let $\Lambda(A) \cdot \Gamma = \Lambda(A_\Gamma)$ where A_Γ is uniquely defined in $\mathfrak{N}$, for every $A \epsilon \mathfrak{N}$, $\Gamma \epsilon \mathrm{P}'$. For every such Γ, A, define $\mathrm{P}_\Gamma A = A_\Gamma$. P_Γ is then in $\Omega(\mathfrak{N})$, and $\{\Lambda_\Gamma = \Gamma, \mathrm{P}_\Gamma, 0, 0\}$ is an $\mathfrak{N} - \mathrm{P}'$-factor set. In the corresponding extension $\mathfrak{P}$ of pairs (Γ, A) every operator on $\mathfrak{N}$ commutative with all the right multipliers of $\mathfrak{N}$ is defined by a left multiplier of $\mathfrak{P}$, $(\Gamma, 0)(0, A) = (0, \Gamma A)$. $\mathfrak{P}$ is called the left holomorph of $\mathfrak{N}$.*

6. Rings of arbitrary type. Let $\mathfrak{E}$ be a ring. $\mathfrak{E}^n$ denotes the set of all products $R_1 \cdots R_n$ of length n, $R_i \epsilon \mathfrak{E}$, $\mathfrak{L}_n$ the left annihilator of $\mathfrak{E}^n$: $\mathfrak{L}_n = \{R \epsilon \mathfrak{E}, R\mathfrak{E}^n = 0\}$, $\mathfrak{R}_n$ the right annihilator of $\mathfrak{E}^n$. These annihilators are ideals, and for rings with suitable chain conditions,[15] one has two series, $0 < \mathfrak{L}_1 < \mathfrak{L}_2 < \cdots < \mathfrak{L}_\lambda = \mathfrak{L}_{\lambda+1} = \cdots$ and $0 < \mathfrak{R}_1 < \mathfrak{R}_2 < \cdots < \mathfrak{R}_\rho = \mathfrak{R}_{\rho+1} = \cdots$. Any ring $\mathfrak{E}$ whose annihilator series repeat at these points is said to be of *type* (ρ, λ), specifically, of *right type* ρ, *left type* λ.

A nilpotent ring of index α has type $(\alpha - 1, \alpha - 1)$, and commutative algebras have type $\rho = \lambda$. Rings with unit are of type $(0, 0)$. Examples of rings with type $\rho \neq \lambda$ are less familiar. A ring with left unit has $\rho = 0$, and may have $\lambda = 1$, e. g., the algebra $e_1{}^2 = e_1$, $e_1e_2 = e_2$, $e_2e_1 = e_2{}^2 = 0$. But rings $\mathfrak{E}$ with left unit always have $\lambda \leqq 1$, for $\mathfrak{E}^2 = \mathfrak{E}$. It is the purpose of this section to indicate a construction for rings of any prescribed type (ρ, λ).

LEMMA. *If $\mathfrak{J}$ is a ring with right annihilator $\mathfrak{R}_1 \neq 0$, and $\mathfrak{J}/\mathfrak{R}_1$ is of right type ρ, then $\mathfrak{J}$ is of right type $\rho + 1$.*

The proof is trivial.

In the following the ring $\mathfrak{C}$ of rational integers $\pm n$ will be used in its ordinary rôle as operator ring on a module $\mathfrak{M} = \{0, A, B, \cdots\}$, i. e., for

---

[15] "Annihilator ideals," *loc. cit.*[5], p. 625.

$n > o, nA = A + \cdots + A$ (n summands); $oA = 0$; for $n < o, nA = (-n)(-A)$.

Lemma. *Let* $\mathfrak{F}_1 = \{o, \sigma, \tau, \cdots\}$ *be of type* $(\rho, 0)$ *and let* $\mathfrak{R}_1$ *be a null ring* $(\mathfrak{R}_1{}^2 = 0)$, *whose additive group admits a ring* Ω *of operators* P_σ *in* $\Omega(\mathfrak{R}_1)$, *such that, for all* A *in* $\mathfrak{R}_1$,

1. $\mathrm{P}_\sigma A + \mathrm{P}_\tau A = \mathrm{P}_{\sigma+\tau} A$.
2. $\mathrm{P}_\sigma(\mathrm{P}_\tau A) = \mathrm{P}_{\tau\sigma} A$.
3. $\mathfrak{F}_1\sigma = o$, *and* $\mathrm{P}_\sigma \mathfrak{R}_1 = 0$ *implies* $\sigma = o$.
4. $\Omega A = 0$ *implies* $A = 0$.
5. $nA = 0$ *implies* $n = o$ *or* $A = 0$, $n \in \mathfrak{C}$.

Then $(0, \mathrm{P}_\sigma, 0, 0)$ *is an* $\mathfrak{R}_1 - \mathfrak{F}_1$*-factor set, and the extension ring* $\mathfrak{F}$ *of pairs* (σ, A) *is of type* $(\rho + 1, 0)$. *Moreover, the null ring* $\mathfrak{R}$ *of pairs* $[m, B]$, $m \in \mathfrak{C}$, $B \in \mathfrak{R}_1$, *with addition* $[m, B] + [n, C] = [m + n, B + C]$ *admits the set* Ω' *of operators* $\mathrm{P}(\sigma, A)$, (σ, A) *in* $\mathfrak{F}$, *defined by* $\mathrm{P}(\sigma, A)[m, B] = [o, \mathrm{P}_\sigma B + mA]$, *and the analogues* 1′-5′ *of* 1-5 *hold in* $\mathfrak{R}$.

Since $\mathfrak{R}_1$ is a null ring, $\mathrm{A}(\mathfrak{R}_1) = 0 = \mathrm{P}(\mathfrak{R}_1)$ and conditions 1, 2 imply that $\sigma \to \mathrm{P}_\sigma$ is an inverse ring homomorphism and that properties 1-10 of § 2 hold for $\{0, \mathrm{P}_\sigma, 0, 0\}$. Multiplication in $\mathfrak{F}$ is given by $(\sigma, A)(\tau, B) = (\sigma\tau, \mathrm{P}_\tau A)$, and 3 implies that $(o, \mathfrak{R}_1)$ is the right annihilator of $\mathfrak{F}$. Then by the preceding Lemma, $\mathfrak{F}$ is of right type $\rho + 1$. If $(\sigma, A)(\mathfrak{F}_1, \mathfrak{R}_1) = (o, 0)$, then $(\sigma, A) = (o, 0)$ using 4, the left type of $\mathfrak{F}_1$ being 0. The properties 1′, 2′, 5′ of the operators $\mathrm{P}(\sigma, A)$ on $\mathfrak{R}$ are readily verified. For 3′, note that $(\mathfrak{F}_1, \mathfrak{R}_1)(\sigma, A) = (o, 0)$ implies $\sigma = o$, and $\mathrm{P}(o, A)[\mathfrak{C}, \mathfrak{R}_1] = [o, \mathfrak{C}A] = [o, 0]$ implies $A = 1A = 0$. Finally, $\mathrm{P}(\mathfrak{F}_1, \mathfrak{R}_1)[n, A] = [o, 0] = [o, \Omega A + n\mathfrak{R}_1]$ implies $\Omega A = 0$ and thus $A = 0, n = o$.

Lemma. *Let* $\mathfrak{F}_1 = \{o, \sigma, \tau, \cdots\}$ *be the ring of rational integers* (*tpye* $0, 0$), $\mathfrak{R}_1 = \{0, A, \cdots\}$ *the null ring with the same additive group as* $\mathfrak{F}_1$. *Define* $\mathrm{P}_\sigma A = \sigma A = A + \cdots + A$ *to* σ *summands* $(\sigma > o)$ *as usual. Then conditions* 1-5 *of the preceding lemma are satisfied.*

The proof is trivial.

The preceding two lemmas thus provide the basis and induction step for the construction of a ring of type $(\rho, 0)$ for arbitrary ρ. The left type is now considered.

Of course the left hand analogues of the preceding lemmas hold and the existence of rings of type $(0, \lambda)$, λ arbitrary, is already established.

Lemma. *Let* $\mathfrak{F}_1 = \{o, \sigma, \tau, \cdots\}$ *be of type* (ρ, λ), $\mathfrak{R}_1$ *a null ring* $(\mathfrak{R}_1{}^2 = 0)$ *whose additive group admits a set* Ω *of operators* Λ_σ *in* $\Omega(\mathfrak{R}_1)$ *such that, for all* A *in* $\mathfrak{R}_1$,

i. $\Lambda_\sigma A + \Lambda_\tau A = \Lambda_{\sigma+\tau} A$.
ii. $\Lambda_\sigma(\Lambda_\tau A) = \Lambda_{\sigma\tau} A$.
iii. $\sigma\mathfrak{F}_1 = 0$ *and* $\Lambda_\sigma \mathfrak{N}_1 = 0$ *implies* $\sigma = o$.
iv. $\Omega^{\rho+1} A = 0$ *implies* $\Omega^\rho A = 0$.[16]
v. $\Omega^\rho \mathfrak{N}_1 \neq 0$.
vi. $nA = 0$ *implies* $n = o$ *or* $A = 0$, $n \in \mathfrak{C}$.

Then $\{\Lambda_\sigma, 0, 0, 0\}$ *is an* $\mathfrak{N}_1 - \mathfrak{F}_1$*-factor set, and the extension ring* $\mathfrak{F}$ *of pairs* (σ, A) *has type* $(\rho, \lambda + 1)$. *Moreover, the null ring* $\mathfrak{N}$ *of pairs* $[m, B]$, $m \in \mathfrak{C}$, $B \in \mathfrak{N}_1$, *with addition* $[m, B] + [n, C] = [m + n, B + C]$ *admits the set* Ω' *of operators* $\Lambda(\sigma, A)$ *in* $\Omega(\mathfrak{N})$ *defined by* $\Lambda(\sigma, A)[n, B] = [o, \Lambda_\sigma B + nA]$, *and the analogues* i′-vi′ *hold in* $\mathfrak{N}$.

The first half of the theorem is readily verified. Note that $\mathfrak{F}^{\rho+1} = (\mathfrak{F}_1{}^{\rho+1}, \Omega^\rho \mathfrak{N}_1)$. Hence $\mathfrak{F}^{\rho+1}(\sigma, A) = (o, 0)$ implies $\mathfrak{F}^\rho(\sigma, A) = (o, 0)$ using iv, and the right type of $\mathfrak{F}_1$. Since there exists in $\mathfrak{F}_1$ a σ such that $\mathfrak{F}_1{}^\rho \sigma = 0$, $\mathfrak{F}_1{}^{\rho-1}\sigma \neq 0$, $(\sigma, 0)$ plays a similar rôle in $\mathfrak{F}$.

Verification of i′-iii′, vi′ is simple. For iv′ note that $\Omega'^{\rho+1}[n, A] = [o, \Omega^{\rho+1} A + n\Omega^\rho \mathfrak{N}_1] = [o, 0]$ implies $\Omega^{\rho+1} A = 0$, thus $\Omega^\rho A = 0$ and $n\Omega^\rho \mathfrak{N}_1 = 0$. Then v, vi imply $n = 0$ and $\Omega'^\rho[n, A] = [o, 0]$.

Finally, from v, there exists a product $\sigma_1 \cdots \sigma_\rho$ and an element A such that $\Lambda_{\sigma_1} \Lambda_{\sigma_2} \cdots \Lambda_{\sigma_\rho} \cdot A \neq 0$. Hence $\Omega'^\rho[\mathfrak{C}, \mathfrak{N}_1]$ contains $\Lambda(\sigma_1 \cdots \sigma_\rho, 0)$. $[o, A]$, and is thus not zero.

Suppose that it is required to construct a ring of arbitrary type (ρ, λ). If either $\rho = 0$ or $\lambda = 0$ the first three lemmas suffice. If not, first construct a ring of type $(\rho, 0)$. It is now clear that to apply the preceding lemma, a basis must be furnished for the inductive construction from a ring of type $(\rho, 0)$ to one of type (ρ, λ). This is done in the

LEMMA. *Let* $\mathfrak{F}_1 = \{o, \sigma, \tau, \cdots\}$ *be a ring of type* $(\rho, 0)$, *with* $\rho \geqq 1$. *Then define* $\mathfrak{N}_1$ *to be the null ring with the same additive group as* $\mathfrak{F}_1$ *and define* $\Lambda_\sigma \cdot \tau = \sigma\tau$, *the left* σ *multiple of* τ *in the ring* $\mathfrak{F}_1$. *Then these operators satisfy the conditions* i-vi *of the preceding lemma.*

The proof is trivial.

THEOREM. *If* ρ, λ *are any non-negative rational integers, there exists a ring of type* (ρ, λ).[17]

UNIVERSITY OF MICHIGAN.

---

[16] See the definition of exponent at the beginning of this section.

[17] This settles a point raised by Professor Saunders Mac Lane in connection with the author's paper "Annihilator Ideals" (*loc. cit.*[5]).

ON SYMMETRIZED KRONECKER POWERS AND THE STRUCTURE OF THE FREE LIE RING.*

By R. M. Thrall.

---

1. **Introduction.** This paper is divided into three chapters. In Chapter I foundations are laid for a general theory of representations of "power type" and their relationship with rings. Kronecker powers, symmetrized Kronecker powers, and the *Lie Representation* are introduced as transformations induced in certain modules of the free non-commutative ring, the free commutative ring, and the free Lie ring, respectively, by a class of ring automorphisms.

In Chapter II the starting point (§ 3) is a general discussion of a problem mentioned by Littlewood:[1] the analysis into irreducible invariant matrices of an invariant matrix of an invariant matrix. This is followed (§ 4) by more specific considerations in the case of the symmetrized Kronecker r-th power of a given invariant matrix. In § 5 formulas are obtained for the anaylsis of the symmetrized Kronecker r-th power of the symmetrized Kronecker m-th power for $r \leqq 3$, all m, and for $m \leqq 2$, all r. The chapter is concluded (§ 6) with a table giving the analysis of the symmetrized Kronecker r-th powers of the irreducible representations of the full linear group defined by partitions (λ) of m for all r, m, (λ) with $rm \leqq 10$.

Let $\boldsymbol{L}$ ($=\boldsymbol{L}_n$) denote the free Lie ring with n generators $r_1, \cdots, r_n$, and let $\boldsymbol{L}^c$ be the module consisting of all forms of degree c in these generators. Under the automorphism $r \rightarrow r' = Ar$ (i. e. $r'_i = \Sigma a_{ij} r_j$) $\boldsymbol{L}^c$ is mapped into itself by a linear transformation whose matrix in some given "coordinate system" we denote by $L_c(A)$. The mapping $A \rightarrow L_c(A)$ is a representation of the full linear group which we call the *c-th Lie Representation.* Many properties of the ring $\boldsymbol{L}$ can be deduced from the analysis of these representations. It is not our purpose in this paper to make such deductions but we mention one of the most important to show the trend of such arguments.

It can be shown that every characteristic ideal $\boldsymbol{I}$ of $\boldsymbol{L}$ can be expressed as the direct sum (in the module sense) of its submodules $\boldsymbol{I}^c = \boldsymbol{I} \frown \boldsymbol{L}^c$ ($\frown$ denotes common part). Each $\boldsymbol{I}^c$ is an invariant subspace of $\boldsymbol{L}^c$ relative to the transformations $L_c(A)$; and, conversely, given any invariant subspace $\boldsymbol{I}^{c*}$ of $\boldsymbol{L}^c$ one can construct at least one characteristic ideal $\boldsymbol{I}$ of $\boldsymbol{L}$ such that $\boldsymbol{I}^c = \boldsymbol{I}^{c*}$.

---

* Received April 18, 1941.

[1] [1] p. 206. If $A \rightarrow T(A)$ is a representation of the full linear group then $T(A)$ is called by Littlewood an *invariant matrix.*

In Chapter III (§ 7) a recursion formula is given from which the analyses of the Lie representations can be made. The proof of the formula is based on E. Witt's extension [2] of a Lie ring to a related associative ring. In (§ 8) this formula is simplified and applied to obtain the analyses into irreducible constituents of the c-th Lie representations for $c \leqq 10$.

CHAPTER I.

2. General rings and representations of power type. Let $\boldsymbol{Z}_n$ be the free distributive but totally non-associative $\boldsymbol{K}$-ring ($\boldsymbol{K}$ any field of characteristic 0) with n generators $z_1, \cdots, z_n$. An ordered set of r elements $w_1, \cdots, w_r$ in $\boldsymbol{Z}_n$ can be combined under multiplication (in the order given) in $N_r = (2r-2)!/r!(r-1)!$ ways.[3] Denote these products by $f_i^r(w_1, \cdots, w_r)$, $i = 1, \cdots, N_r$. Since $\boldsymbol{Z}_n$ is distributive any element in it can be written as a linear $\boldsymbol{K}$-combination of monomials, i. e. products of the generators. Since $\boldsymbol{Z}_n$ is free the $N_r n^r$ distinct monomials $f_i^r(j) = f_i^r(z_{j_1}, \cdots, z_{j_r})$ are linearly independent and constitute a basis for the $\boldsymbol{K}$-module $\boldsymbol{Z}_n^r$ consisting of all forms of degree r in the generators of $\boldsymbol{Z}_n$.

With the customary definitions of *homomorphism* and *ideal* the usual homomorphism theorems hold good. An ideal $\boldsymbol{J}$ in $\boldsymbol{Z}_n$ is said to be *homogeneous* if

$$\sum_{i,r,(j)} a_i^r(j) f_i^r(j) \in \boldsymbol{J} \text{ requires } \sum_{i,(j)} a_i^r(j) f_i^r(j) \in \boldsymbol{J}, \qquad r = 1, 2, \cdots.$$

Let $A = \| a_{ik} \|$ be any non-singular n-rowed $\boldsymbol{K}$-matrix. Then $z_i \to z'_i = \Sigma a_{ik} z_k$ is an automorphism of $\boldsymbol{Z}_n$. In particular the $\boldsymbol{K}$-modules $\boldsymbol{Z}_n^r$ are invariant under A. Denote by $f_i^r(j)'$ the product $f_i^r(j)$ with z'_j substituted for z_j. Then the $N_r n^r$-rowed matrix $A[r] = \| a^{[r]}_{(j)'(j)} \|$ whose rows are the coefficients in the expressions for the $f_i^r(j)'$ in terms of the $f_i^r(j)$ completely describes the effect in $\boldsymbol{Z}_n^r$ of the application of A to $\boldsymbol{Z}_n$. It is easy to verify that $AB = C$ implies $A[r]B[r] = C[r]$. Hence if $\boldsymbol{A}$ is any set of n-rowed matrices A, closed under multiplication, then $A \to A[r]$ is a representation $\boldsymbol{A}[r]$ of $\boldsymbol{A}$. In particular $A[1] = A$.

Let $\boldsymbol{J}$ be any homogeneous ideal in $\boldsymbol{Z}_n$ which is invariant under every A in $\boldsymbol{A}$, and denote by $\boldsymbol{R}$ the residue class ring $\boldsymbol{Z}_n - \boldsymbol{J}$. In the homomorphism between $\boldsymbol{Z}_n$ and $\boldsymbol{R}$ let $z_i \to x_i$. Then the mapping $x_i \to x'_i = \Sigma a_{ik} x_k$ is an automorphism of $\boldsymbol{R}$ (provided $\det A \neq 0$). The homogeneity of $\boldsymbol{J}$ implies that $\boldsymbol{Z}_n^r$ is mapped into the $\boldsymbol{K}$-module $\boldsymbol{R}^r$ generated (under $+$) by all products of degree r in the x_i. Hence $\mathbf{R}^r$ is invariant under A. Proceeding as above, let $A(r)$ be the matrix which describes the transformation in $\mathbf{R}^r$ induced by A, in terms of some set of basis elements for $\mathbf{R}^r$. Then the mapping $A \rightarrow A(r)$ constitutes a representation $\mathbf{A}(r)$ of $\mathbf{A}$. The elements of $A(r)$ are forms of degree r in the elements of A. We call any representation $\mathbf{A}(r)$ a *representation of power type of* $\mathbf{A}$.

---

[2] [7] Theorem I, p. 152.

[3] [6] p. 121.

If an ideal $\mathbf{J}$ is invariant under the full linear group, $GL(n)$, it is, *a fortiori*, invariant under any closed system $\mathbf{A}$. In the following chapters we shall consider three choices for the ideal $\mathbf{J}$, all invariant under $GL(n)$.

1. *The free associative ring* $\mathbf{X}_n$. Let $\mathbf{J}$ contain all elements $a(bc) - (ab)c$ as $a, b, c,$ range through $\mathbf{Z}_n$. Considered as a $\mathbf{K}$-module $\mathbf{J}$ has the basis $f_i{}^r(j) - f_i{}^r(j)$ for all choices of $i, r, (j)$, and so is homogeneous. We set $\mathbf{Z}_n - \mathbf{J} = \mathbf{X}_n$ and denote by x_i the image of z_i in the homomorphism of $\mathbf{Z}_n$ on $\mathbf{X}_n$. The representation $A \rightarrow A(r)$ is in this case the ordinary Kronecker r-th power, and we shall denote $A(r)$ by $\Pi_r(A)$.

It is of some interest to note that by a proper ordering of the basis elements in $\mathbf{Z}_n{}^r$ the matrix $A[r]$ becomes $\Pi_r(A) \times E_{N_r}$ ($\times$ denoting Kronecker product). For the elements $f_i{}^r(j)$, with i and r fixed, are by virtue of the distributive law transformed among themselves, in the $\mathbf{K}$-module which they generate, in precisely the same manner as the monomials in $\mathbf{X}_n{}^r$.

II. *The free associative and commutative ring* $\mathbf{Y}_n$. Add to the set of generators for $\mathbf{J}$ in I the set of all elements $ab - ba$; and denote the resulting residue class ring by $\mathbf{Y}_n$. $\mathbf{Y}_n$ is the free commutative polynomial ring with n generators y_i (the images of the z_i). In this case we denote $A(r)$ by $S_r(A)$. The representation $A \rightarrow S_r(A)$ is the *symmetrized Kronecker r-th power*. The degree of $S_r(A)$ is $\binom{n+r-1}{r}$.

III. *The free Lie ring*, $\mathbf{L}_n$. Let $\mathbf{J}$ be generated by the elements aa and $a(bc) + c(ab) + b(ca)$ as a, b, c run through $\mathbf{Z}_n$. Then $\mathbf{Z}_n - \mathbf{J}$ is the free Lie ring $\mathbf{L}_n$ with n generators $s_1, \cdots, s_n$. In this case we denote $A(r)$ by $L_r(A) = ||a_{jk}^{(r)}||$. This is the *r-th Lie representation* mentioned above. We propose to study its properties, giving, in particular, its relation to the representations $\Pi_r(A)$ and $S_r(A)$.

CHAPTER II.

3. On the analysis of an invariant matrix of an invariant matrix. Let $B = T(A)$ be an N rowed representation, $\mathbf{T}$, of $GL(n)$ and let $C = T^*(B)$ be a representation, $\mathbf{T}^*$, of $GL(N)$. Then $A \rightarrow C = T^*(T(A))$ is a representation of $GL(n)$. In Littlewood's language[4] C is an invariant matrix of an invariant matrix. In this section we discuss the analysis of C in case both $\boldsymbol{T}$ and $\boldsymbol{T}^*$ are irreducible; i. e. we set $B = A(\lambda)$, $C = B(\mu) = A(\lambda)(\mu)$ where (λ) and (μ) are partitions of m and r respectively and $A(\lambda)$ denotes the irreducible representation of $GL(n)$ defined by (λ).[5]

Denote the eigenvalues[6] of A by $(z) = (z_1, \cdots, z_n)$ and set $s_i = \operatorname{tr} A^i = z_1{}^i + \cdots + z_n{}^i$. We shall use Murnaghan's notation[7] $q_m(s) = q_m(s_1, \cdots, s_m) = p_m(z) = p_m(z_1, \cdots, z_n)$ for the complete symmetric functions of the z's. Let $\zeta_1, \zeta_2, \cdots$ be the eigenvalues of B and set $t_i = \operatorname{tr} B^i = \zeta_1{}^i + \zeta_2{}^i + \cdots$. We set $q_m(s, i) = q_m(s_i, s_{2i}, \cdots, s_{mi}) = p_m(z^i) = p_m(z_1{}^i, \cdots z_n{}^i)$. Now applying the formulas for the traces of C and B we have[8]

$$\operatorname{tr} C = \sum_{(\beta)} \frac{\chi^{(\beta)}_{(\mu)}}{\beta_1! \cdots \beta_r!} \left(\frac{t_1}{1}\right)^{\beta_1} \cdot \cdots \cdot \left(\frac{t_r}{r}\right)^{\beta_r} \tag{3.1}$$

and

$$t_i = \sum_{(\alpha)} \frac{\chi^{(\alpha)}_{(\lambda)}}{\alpha_1! \cdots \alpha_m!} \left(\frac{s_i}{1}\right)^{\alpha_1} \cdot \cdots \cdot \left(\frac{s_{mi}}{m}\right)^{\alpha_m}. \tag{3.2}$$

Substituting 3.2 in 3.1 we obtain $\operatorname{tr} C$ as an isobaric polynomial of weight rm in $s_1, \cdots, s_{rm}$, say

$$\operatorname{tr} C = Q_{(\lambda)(\mu)}(s) \tag{3.3}$$

Set $s^{(\nu)} = s_1{}^{\nu_1} \cdot \cdots \cdot s_M{}^{\nu_M}$, $\nu_1 + \cdots + M\nu_M = M$. Then[9]

$$s^{(\nu)} = \Sigma \chi^{(\nu)}_{(v)} \{v\} \tag{3.4}$$

where $\{v\}$ denotes $\operatorname{tr} A(v)$ and the sum extends over all partitions (v) of M.

Now take $M = rm$ in 3.4, substitute in 3.3, and we obtain for $\operatorname{tr} C$ an expression as a sum of irreducible characters, $\{v\}$, which of course determines uniquely the irreducible constituents of the representation $A \rightarrow C$.

This method depends upon the availability of the character tables for the symmetric groups of degree m, r, and rm. From the product decomposition rule[10] for the irreducible characters $\{v\}$ and the formula[11] $s_i = \{i\} - \{i-1, 1\} + \cdots + (-1)^i\{1^i\}$ one can obtain the decomposition of $\operatorname{tr} C$ from 3.3

---

[4] [1] p. 206.

[5] E. g. Murnaghan's $M_{(\lambda)}(A)$, [3] p. 129.

[6] These eigenvalues lie in some algebraic extension of $\boldsymbol{K}$. However, the final results obtained are rational and therefore valid in $\boldsymbol{K}$.

[7] [3] Chapter IV.

[8] [3] p. 133. The χ's are irreducible characters of the symmetric groups of degree r and m.

[9] [1] p. 86.

[10] [5] p. 750.

[11] [3] p. 108.

without reference to the characters of the symmetric group of degree rm. Then applying the formula[12] $\operatorname{tr} A(\lambda) = \det \| q_{\lambda_i+j-i}(s) \|$ and the similar formula for $\operatorname{tr} B(\mu)$ one can avoid altogether the use of characters of the symmetric group. So far as actual computation is concerned this method is little quicker than first computing all characters needed and then applying the first process. A modification of this method which is used below depends upon the fact that products of the functions $q_m(s)$ are somewhat easier to analyze than products of the power sums s_i. For $q_m(s) = p_m(z) = \{m\}$. To apply this fact we must replace 3. 3 by an expression for $\operatorname{tr} C$ as an (isobaric) polynomial, $P_{(\lambda)(\mu)}(p_1, \cdots, p_{rm})$, (the arguments (z) being omitted from $p_m(z)$). Using the formula $t_i = \det \| q_{\lambda_j+k-j}(s, i) \|$ this problem can be solved as soon as we have an expression for $q_m(s, k) = p_m(z^k)$ in terms of the p's.

Denote by $T_{(\lambda)}(z)$ the ordinary (power sum) symmetric function $\Sigma z_1^{\lambda_1} \cdot \cdots \cdot z_k^{\lambda_k}$.

THEOREM I. $p_m(z^k) = \Sigma p_{\lambda_1} \cdot \cdots \cdot p_{\lambda_k} T_{(\lambda)}(\omega)$ *the sum extending over all partitions* $(\lambda) = (\lambda_1, \cdots, \lambda_k)$ *of* mk *such that* $mk \geqq \lambda_1 \geqq \cdots \geqq \lambda_k \geqq 0$ *and with* $\omega_\nu = \exp(2\pi i\nu/k)$.

We recall that[13]

$$(3.5) \quad 1/\det(E - \tau A) = \prod_{\beta=1}^{\beta=m} (1 - z_\beta \tau)^{-1} = p_0 + p_1\tau + \cdots + p_\nu\tau^\nu + \cdots.$$

Replace A by A^k and τ by τ^k in 3. 5 and we get

$$(3.6) \quad p_0 + p_1(z^k)\tau^k + p_2(z^k)\tau^{2k} + \cdots = \prod_{\beta=1}^{\beta=m} (1 - z_\beta{}^k\tau^k)^{-1}$$

$$= \prod_{\beta=1}^{\beta=m} \prod_{\gamma=1}^{\gamma=k} (1 - \omega_\nu z_\beta \tau)^{-1} = \prod_{\gamma=1}^{\gamma=k} [p_0 + p_1\omega_\gamma\tau + p_2\omega_\gamma{}^2\tau^2 + \cdots].$$

The theorem now follows upon equating coefficients of τ^{mk} in the first and last terms of 3. 6.

In particular

$$s_m = p_1(z^m) = \Sigma p_{\lambda_1} \cdot \cdots \cdot p_{\lambda_m} T_{(\lambda)}(\omega)$$

the sum extending over all partitions (λ) of m. We have also the formula:[14]

$$s_m = \Sigma (-1)^{\alpha_0-1} m(\alpha_0 - 1)!\, p_1{}^{\alpha_1} \cdot \cdots \cdot p_m{}^{\alpha_m} / \alpha_1! \cdot \cdots \cdot \alpha_m!$$

where the sum extends over all partitions $(\alpha): \alpha_1 + 2\alpha_2 + \cdots + m\alpha_m = m$ of m and $\alpha_0 = \alpha_1 + \cdots + \alpha_m$. Equating coefficients of the power products of the p's we get as a side result of our theorem:

---

[12] [3] p. 125.

[13] [3] p. 108.

[14] [2] p. 7.

COROLLARY I. *$T_{(\lambda)}(\omega) = (-1)^{\alpha_0-1}m(\alpha_0-1)!/\alpha_1!\cdots\cdots\alpha_m!$ where α_i is the number of λ's equal to i $(i > 0)$ and α_0 is the number of non-zero λ's.*

We close this section with a theorem which will cut in half the labor of analyzing all representations $A \to C = A(\lambda)(\mu)$ for given r and m. We denote by $(\bar{\lambda})$ and $(\bar{\mu})$ the partitions of m and r associated[15] with (λ) and (μ) respectively. If $A \to T(A)$ is any matrix representation of $GL(n)$ whose elements are forms of degree M in the elements of A and whose trace is $\operatorname{tr} T(A) = Q(s_1, s_2, \cdots, s_M)$ we denote by $A \to \bar{T}(A)$ the representation[16] of $GL(n)$ defined by $\operatorname{tr} \bar{T}(A) = Q(s_1, -s_2, \cdots, (-1)^{M-1}s_M)$, and call the two representations *associated.* (Thus the irreducible representation defined by (λ) is associated with the representation defined by $(\bar{\lambda})$.[17])

THEOREM II. *Set $C = A(\lambda)(\mu)$ and $D = A(\bar{\lambda})(\mu)$. Then if m is even $\bar{D} = C$ and if m is odd $\bar{D} = A(\lambda)(\bar{\mu})$.*

Proof. Set $\bar{t}_i = \operatorname{tr} \bar{B}^i$ where $B = A(\lambda)$. Substitute

$$\chi^{(\alpha)}_{(\bar{\lambda})} = (-1)^{\alpha_2+\alpha_4+\cdots}\chi^{(\alpha)}_{(\lambda)}$$

in 3.2, apply the fact that $m - \alpha_1 - \alpha_3 - \cdots$ is even and we obtain

$$\begin{aligned} \bar{t}_i &= \sum_{(\alpha)} \frac{\chi^{(\alpha)}_{(\lambda)}}{\alpha_1!\cdots\alpha_m!}\left(\frac{s_i}{1}\right)^{\alpha_1}\left(\frac{-s_{2i}}{2}\right)^{\alpha_2}\cdots \\ &= (-1)^m \sum_{(\alpha)} \frac{\chi^{(\alpha)}_{(\lambda)}}{\alpha_1!\cdots\alpha_m!}\left(\frac{-s_i}{1}\right)^{\alpha_1}\left(\frac{-s_{2i}}{2}\right)^{\alpha_2}\cdots \end{aligned}$$

Use the first form if i is odd and the second if i is even and we get $\bar{t}_i = (-1)^{m(i-1)}\tilde{t}_i$ where $\tilde{t}_i = t_i(s_1, -s_2, \cdots)$. Now apply 3.1 to obtain

$$\begin{aligned} \operatorname{tr} D &= \sum_{(\beta)} \frac{\chi^{(\beta)}_{(\mu)}}{\beta_1!\cdots\beta_r!}\left(\frac{\bar{t}_1}{1}\right)^{\beta_1}\left(\frac{\bar{t}_2}{2}\right)^{\beta_2}\cdots \\ &= \sum_{(\beta)} \frac{\chi^{(\beta)}_{(\mu)}}{\beta_1!\cdots\beta_r!}\left(\frac{\tilde{t}_1}{1}\right)^{\beta_1}\left(\frac{(-1)_m\tilde{t}_2}{2}\right)^{\beta_2}\cdots \end{aligned}$$

If m is odd this is equal to

$$\sum_{(\beta)} \frac{\chi^{(\beta)}_{(\bar{\mu})}}{\beta_1!\cdots\beta_r!}\left(\frac{\tilde{t}_1}{1}\right)^{\beta_1}\left(\frac{\tilde{t}_2}{2}\right)^{\beta_2}\cdots = \operatorname{tr} \tilde{E}$$

where $E = A(\lambda)(\bar{\mu})$, and if m is even it is

$$\sum_{(\beta)} \frac{\chi^{(\beta)}_{(\mu)}}{\beta_1!\cdots\beta_r!}\left(\frac{\tilde{t}_1}{1}\right)^{\beta_1}\left(\frac{\tilde{t}_2}{2}\right)^{\beta_2} = \operatorname{tr} \tilde{C}$$

which completes the proof of the theorem.

---

[15] [3] p. 120.

[16] Throughout this paper we consider equivalent representations as equal.

[17] [3] p. 120.

4. The symmetrized Kronecker r-th power of a reducible representation. If $A \rightarrow B = T(A)$ is any representation of $GL(n)$ and $C = S_r(B)$ then[18] $\operatorname{tr} C = q_r(t)$ where $t_i = \operatorname{tr} B^i$. In case $A \rightarrow B$ is an irreducible representation we have the situation treated in § 3 for the partition $(\mu) : r = r$ of r. We now derive a formula which enables us to reduce the analysis of the symmetrized Kronecker r-th power of any representation to the analysis of Kronecker products and of symmetrized powers of irreducible representations.

Suppose that $B = B_1 + B_2 + \cdots$ where the B_ν are any representations, reducible or irreducible, and not necessarily distinct. Set $\operatorname{tr} B_\nu{}^i = t(\nu, i)$ and denote $(t(\nu,1), t(\nu,2), \cdots, t(\nu,r))$ by $(t(\nu))$. Then obviously $t_i = t(1,i) + t(2,i) + \cdots$, which is for $i = 1$ the case $r = 1$ of the following lemma.

LEMMA I. $q_r(t) = \Sigma q_{r_1}(t(1)) q_{r_2}(t(2)) \cdots$ *where the sum extends over all non-negative integral solutions of the equations* $r_1 + r_2 + \cdots = r$.

(Note that the number of terms in the summation is the same as the number of distinct monomials in the expansion of $(y_1 + y_2 + \cdots)^r$.)

Proof by induction on r. $q_r(t) = q_r(t(1) + t(2) + \cdots)$ is isobaric of weight r in the set of traces $t(\nu)$. Hence to make an induction proof it is sufficient to show that differentiation of each side with respect to each $t(\nu, i)$ gives a true relation (i. e. true relative to our induction hypothesis). On the left hand side we get[19]

$$p\delta q_r(t)/\delta t(\nu, p) = q_{r-p}(t) \tag{4.1}$$

Denote the right hand side by $Q_r(t)$. Then since a "q" with a negative subscript is zero we have

$$p\delta Q_r(t)/\delta t(\nu, p) = \Sigma q_{r_1}(t(1)) q_{r_2}(t(2)) \cdots q_{r_\nu - p}(t(\nu)) \cdots = Q_{r-p}(t) \tag{4.2}$$

for the summation is over all non-negative integral solutions of the equation $r_1 + r_2 + \cdots + (r_\nu - p) + \cdots = r - p$.

This establishes the lemma. To justify the statement preceding the lemma we require that each B_ν be irreducible and observe that $q_{r_\nu}(t(\nu)) = \operatorname{tr} S_{r_\nu}(B_\nu)$. Observe that a multiple of an irreducible representation is here treated like any reducible representation. For instance, $S_2(2B) = 2S_2(B) + B \times B$. Strictly speaking the above proof of the lemma holds good only if the constituents B_ν are distinct. However, if several of them are equal one can first polarize and then proceed as before.

5. Symmetrized powers of symmetrized powers. The general problem is the analysis of the representation $A \rightarrow C = S_r(B)$ where $B = S_m(A)$. In this section we give formulas for this analysis for all m if $r \leqq 3$ and for all r if $m \leqq 2$. These formulas and their proofs bear a certain resemblance but the road to a formula for arbitrary r and m is not yet clear. If either m or r is 1 then C is irreducible and there is no problem. We consider first the case $m = 2$, r arbitrary.

---

[18] [3] p. 90.

[19] [3] p. 107.

THEOREM III. *Let $B = S_2(A)$ and $C = S_r(B)$. Then*

$$\operatorname{tr} C = q_r(t) = \Sigma\{2\lambda_1, \cdots, 2\lambda_\kappa\} \tag{5.1}$$

where the sum is over all partitions (λ) *of* r.

Proof by induction on r, and by differentiation. Denote the right hand side of 5.1 by $Q_r(s)$. We establish the theorem by showing that $q_r(t)$ and $Q_r(s)$ satisfy the same differential equations. Since $t_i = (s_i^2 - s_{2i})/2$ we have

$$p\delta q_r(t)/\delta s_p = s_p q_{r-p}(t) + q_{r-p/2}(t) \tag{5.2}$$

where we understand that the second term on the right occurs only when p is even and the first only for $p \leqq r$. We now show that 5.2 is satisfied by $Q_r(s)$.

We have[20]

$$p\delta Q_r(s)/\delta s_p = \sum_{(\lambda)} \sum_{j=1}^{j=\kappa} \{2\lambda_1, \cdots, 2\lambda_j - p, \cdots, 2\lambda_\kappa\} \tag{5.3}$$

where the ordering relation[21] $\{\cdots a, b \cdots\} = -\{\cdots b-1, a+1 \cdots\}$ is to be applied (several times if necessary) in case $2\lambda_j - p < 2\lambda_{j+1}$.

It is not hard to show[22] that

$$s_p\{\lambda\} = \sum_{j=1}^{j=\kappa+1} \{\lambda_1, \cdots, \lambda_j + p, \cdots, \lambda_\kappa\} \qquad (\lambda_{\kappa+1} = 0),$$

and therefore that

$$s_p Q_{r-p}(s) = \sum_{(\lambda)} \sum_{j=1}^{j=\kappa+1} \{2\lambda_1, \cdots, 2\lambda_j + p, \cdots, 2\lambda_\kappa\} \tag{5.4}$$

where again disordered partitions are to be rearranged and the sum is this time over all partitions (λ) of $r - p$.

If a term in 5.3 is disordered, the ordered term obtained from it is either zero or of the form

$$\{u_1, \cdots, u_j, u_{j+1}, \cdots, u_{j+k}, \cdots, u_l\} \tag{5.5}$$

where $u_{j+1}, \cdots, u_{j+k}$ are odd, the other u's are even, and the sum of the u's is $2r - p$.

---

[20] [3] p. 133.

[21] [4] p. 214.

[22] Cf. [3] p. 165 or [1] p. 89.

We look at the problem in reverse and ask, for a given partition (u) of $2r-p$ of the form in 5. 5, how many times does $\{u\}$ occur in 5. 3 and 5. 4 (i. e. after the terms are ordered and taking account of signs)?

First suppose $k>0$. If $\{u\}$ comes from a term in 5. 3 we can apply the ordering formula in reverse to obtain an "equivalent" partition with at most one odd term and at most one point of disordering; and it must be possible to remove that disordering by adding p to some term. Subject to these restrictions we can reduce a partition of the form of 5. 5 in at most two ways: we can move either u_{j+k+1} or u_{j+k} to the left until u_j is reached. These give, respectively,

$$(5.6)\quad (-1)^k\{u\}=\{\cdots u_j, u_{j+k+1}-k, u_{j+1}+1,\cdots, u_{j+k}+1, u_{j+k+2},\cdots\}$$

and

(5.7)

$$(-1)^{k-1}\{u\}=\{\cdots u_j, u_{j+k}-k+1, u_{j+1}+1,\cdots, u_{j+k-1}+1, u_{j+k+1},\cdots\}.$$

Since adding p to the term $u_{j+k}-k+1$ is to restore order we have

Lemma II. *The right hand side of* 5. 6 *occurs in* 5. 3 *if and only if*

$$\text{(A)}\qquad u_j \geqq u_{j+k+1}-k+p \geqq u_{j+1}+1$$

and the right hand side of 5. 7 *occurs in* 5. 3 *if and only if*

$$\text{(B)}\qquad u_j \geqq u_{j+k}-k+p+1 \geqq u_{j+1}+1.$$

A similar argument applied to 5. 4 gives

$$(5.8)\quad (-1)^k\{u\}=\{\cdots u_{j-1}, u_{j+1}-1,\cdots, u_{j+k}-1, u_j+k, u_{j+k+1},\cdots\}$$

(5. 9)

$$(-1)^{k-1}\{u\}=\{\cdots u_j, u_{j+2}-1,\cdots, u_{j+k}-1, u_{j+1}+k-1, u_{j+k+1},\cdots\}$$

with the corresponding conclusion

Lemma III. *The right hand side of* 5. 8 *occurs in* 5. 4 *if and only if*

$$\text{(C)}\qquad u_{j+k}-1 \geqq u_j+k-p \geqq u_{j+k+1}$$

and the right hand side of 5. 9 *occurs in* 5. 4 *if and only if*

$$\text{(D)}\qquad u_{j+k}-1 \geqq u_{j+1}+k-p-1 \geqq u_{j+k+1}.$$

If we take into account the signs with which $\{u\}$ occurs in 5. 6, $\cdots$, 5. 9 the following lemma shows that the difference of the right hand sides of 5. 3 and 5. 4, when its terms are ordered, contains no terms $\{u\}$ with $k>0$.

Lemma IV. i) A *and* D *cannot both hold*; ii) B *and* C *cannot both hold*; iii) *the truth of either* A *or* D *is equivalent to the truth of either* B *or* C.

Proof. i) If A and D were both true then from $u_{j+k+1} \geqq u_{j+1} + 1$ and $u_{j+1} + k - p - 1 > u_{j+k+1}$ we obtain by addition $-1 \geqq +1$. ii) follows similarly.

We establish just one of the four parts of iii); the others follow similarly. Suppose then that A is true and both B and C are false. Since (u) is an ordered partition with u_{j+k} odd and u_{j+k+1} even we have $u_{j+k} - k + p + 1 \geqq u_{j+k+1} - k + p + 2$ and so A true, B false requires $u_{j+k} - k + p + 1 > u_j$ and in the same fashion we show that A true, C false requires $u_j + k - p > u_{j+k} - 1$. This gives $u_j + 2 > u_{j+k} - k + p + 1 > u_j$ which is impossible for even numbers. (Note that p and k have like parity.)

There remains the case $k = 0$, which requires p even, and, of course, the terms $\{u\}$ of this kind must be already in order as they stand in 5.3 and 5.4 (for application of the ordering process introduces odd terms). If $u_{j-1} - p \geqq u_j$ for any j between 2 and $l + 1$ $(u_{l+1} = 0)$ then the partition $(\cdots, u_j + p, \cdots)$ gives rise to a term $\{u\}$ in 5.3 and the partition $(\cdots, u_{j-1} - p, \cdots)$ gives rise to a term $\{u\}$ in 5.4. The partition $(u_1 + p, \cdots)$ gives a term $\{u\}$ in 5.3 not matched by any term from 5.4. Hence $\{u\}$ occurs once more in 5.3 than in 5.4. But this is just compensated by adding $Q_{r-p/2}(s)$ to 5.4 and thus completes the proof of the formula

$$p\delta Q_r(s)/\delta s_p = s_p Q_{r-p}(s) + Q_{r-p/2}(s). \tag{5.10}$$

Theorem III now follows at once by our induction hypothesis.

Applying Theorem II with $m = 2$ we have:

COROLLARY II. *Let $B = A(1^2)$ and $C = S_r(B)$. Then*

$$\operatorname{tr} C = \Sigma \{\lambda_1, \lambda_1, \lambda_2, \lambda_2, \cdots, \lambda_\kappa, \lambda_\kappa\}$$

where the sum is over all partitions (λ) of r.

Next consider the problem of this section for m arbitrary and $r = 2$.

THEOREM IV. *Let $C = S_2(B)$ where $B = S_m(A)$. Then*

$$\operatorname{tr} C = \Sigma \{2m - 2v, 2v\}$$

where the sum is over all integers v such that $2v \leqq m$.

Proof. By 3.1 we have $\operatorname{tr} C = (t_1^2 + t_2)/2$ where $t_i = q_m(s, i)$. Now $t_1^2 = \{m\}\{m\} = \{2m\} + \{2m-1, 1\} + \cdots + \{m^2\}$. To calculate t_2 we apply Theorem I with $k = 2$. Observe that for $(\lambda) = (2m - j, j)$ we have $T_{(\lambda)}(\omega) = 2(-1)^j$ for $j < m$ and $= (-1)^m$ for $j = m$. Hence $t_2 = 2\{2m\} - 2\{2m-1\}\{1\} + 2\{2m-2\}\{2\} - \cdots + (-1)^m\{m\}\{m\}$. Simplify this by the multiplication formula $\{2m - j\}\{j\} = \{2m\} + \{2m-1, 1\} + \cdots$

$+ \{2m - j, j\}$. Then the coefficient of $\{2m - v, v\}$ in the resulting expression for t_2 is $2(-1)^v + 2(-1)^{v+1} + \cdots + (-1)^m = (-1)^v$ and so $t_2 = \{2m\} - \{2m-1, 1\} + \cdots + (-1)^m\{m^2\}$ from which the theorem follows at once.

By application of Theorem II one can obtain a similar formula for the symmetrized square of $B = A(1^m)$; here, of course, the cases m even and m odd require separate attention.

We conclude this section with the analysis of the representation $A \to C = A(m)(3)$. A fundamental tool in handling it is the following formula which gives the analysis into irreducible constituents of the Kronecker product of three symmetrized Kronecker power representations of the full linear group.

LEMMA V. *Let* $M = u_1 + u_2 + u_3$, *where* $u_1 \geqq u_2 \geqq u_3$. *Then*

$$\{u_1\}\{u_2\}\{u_3\} = \Sigma k(u)(\lambda)\{\lambda\}$$

the sum extending over all partitions (λ) *of* M *into at most three parts, where* $k(u)(\lambda) = \max(0, 1 + \min(u_2, \lambda_2, u_1 + u_2 - \lambda_2) - \max(\lambda_3, u_1 + u_2 - \lambda_1))$.

This lemma follows by a direct computation using G. de B. Robinson's rule[23] for multiplying characters of $GL(n)$.

If $u_1 = u_2 = u_3 = m$ we shall abbreviate $k(u)(\lambda)$ to $k(\lambda)$ and observe that

$$(5.11) \quad \begin{aligned} &k(\lambda) = \min(1 + \lambda_1 - \lambda_2, 1 + \lambda_2 - \lambda_3) \ \textit{or equivalently} \\ &k(\lambda) = 1 + \lambda_1 - \lambda_2 \ \textit{if} \ \lambda_2 \geqq m, = 1 + \lambda_2 - \lambda_3 \ \textit{if} \ \lambda_2 \leqq m. \end{aligned}$$

Set $g(\lambda) = 1$, 0, or -1 according as $k(\lambda)$ is congruent to 1, 0, or -1 mod 3.

LEMMA VI. $t_3 = q_m(s, 3) = \Sigma g(\lambda)\{\lambda\}$, *the summation extending over all partitions* (λ) *of* $3m$ *into at most three parts.*

Proof by induction on m *and the differentiation argument.* We denote the summation on the right hand side by $Q_m(s)$ and show that the relations

$$(5.12)' \quad p\delta q_m(s,3)/\delta s_{3p} = q_{m-p}(s,3)\,; \ \delta q_m(s,3)/\delta s_p = 0 \ \text{if} \ p \not\equiv 0 \bmod 3$$

remain true with $Q_m(s)$ substituted for $q_m(s,3)$.

Starting with

$$p\delta Q_m(s)/\delta s_p = \Sigma g(\lambda)\,[\{\lambda_1 - p, \lambda_2, \lambda_3\} + \{\lambda_1, \lambda_2, -p, \lambda_3\} + \{\lambda_1, \lambda_2, \lambda_3 - p\}]$$

we follow the procedure of Theorem III: i. e. reorder disordered partitions, and then calculate from the resulting expression the coefficient of $\{u\}$ for all partitions (u) of $3m - p$ into at most three parts. To establish the lemma we have to show that if p is not divisible by 3 the coefficient of $\{u\}$ is zero and that if p is divisible by 3 the coefficient of $\{u\}$ is three times its coefficient, $g(u)$, in the formula for $Q_{m-p/3}(s)$.

---

[23] [5] p. 750.

To do this we turn the problem around and determine the partitions (λ) of $3m$ for which $\{\lambda\}$ can be obtained, either directly or after restoring order, from $\{u\}$ after adding p to each of its terms in turn. These $\{\lambda\}$ are

$$(5.13)\quad \begin{array}{ll} \text{a)} & \{u_1+p, u_2, u_3\} \\ \text{b)} & \{u_1, u_2+p, u_3\} \text{ or } -\{u_2+p-1, u_1+1, u_3\} \\ \text{c)} & \{u_1, u_2, u_3+p\} \text{ or } -\{u_1, u_3+p-1, u_2+1\} \\ & \qquad\qquad\qquad\quad \text{ or } \{u_3+p-2, u_1+1, u_2+1\} \end{array}$$

We now observe that for any partition (λ) of $3m$, $1+\lambda_1-\lambda_2 \equiv 1+\lambda_2-\lambda_3 \bmod 3$ and so we may always use, say, $1+\lambda_2-\lambda_3$ to determine $g(\lambda)$. Observe also that, computing formally, $g(\lambda_1, \lambda_2, \lambda_3) = -g(\lambda_2-1, \lambda_1+1, \lambda_3) = -g(\lambda_1, \lambda_3-1, \lambda_2+1)$ and so it is sufficient to consider only the first alternatives in b) and c) of 5.13.

Proceeding thus we get: from a) $g(\lambda) \equiv 1+u_2-u_3$; from b) $g(\lambda) \equiv 1+u_2-u_3+p$; and from c) $g(\lambda) \equiv 1+u_2-u_3-p$. Now if p is divisible by 3 these g's are all equal to each other and their sum is $3g(u)$; if p is not divisible by 3 the g's are incongruent mod 3 and therefore have the sum zero. This completes the proof of Lemma VI

LEMMA VII. *Let* $t_i = q_m(s, i)$. *Then*

$$t_1 t_2 = \Sigma h(\lambda)\{\lambda\}$$

where $h(\lambda) = 0$ *if* $k(\lambda)$ *is even and* $h(\lambda) = (-1)^{\lambda_2}$ *if* $k(\lambda)$ *is odd.*

Using the value obtained above for t_2 the proof follows by direct computation. We omit the details.

Now from Lemmas V, VI, and VII we get:

THEOREM V. *Let* $C = A(m)(3)$. *Then*

$$\operatorname{tr} C = q_3(t) = \Sigma f(\lambda)\{\lambda\}$$

where the summation extends over all partitions (λ) *of* $3m$ *into not more than three parts, and*

$$f(\lambda) = (k(\lambda) + 3h(\lambda) + 2g(\lambda))/6.$$

The following verbal description of $f(\lambda)$ is perhaps easier to apply: *To* $k(\lambda) = \min(1+\lambda_1-\lambda_2, 1+\lambda_2-\lambda_3)$ *we add whichever one of* $-2, 0, +2$ *will give a result divisible by* 3. *If this result is even divide by* 6 *to get* $f(\lambda)$. *If this result is odd add or subtract* 3 *according as* λ_2 *is even or odd and then divide by* 6 *to get* $f(\lambda)$.

6. Analysis of certain symmetrized Kronecker powers. Set $\chi_r(\lambda) = \operatorname{tr} C$ where $C = S_r(B)$, $B = A(\lambda)$, and (λ) is a partition of m. If $M = rm$ is prime the representation $A \to C$ is irreducible and there is no problem. The cases with $m = 2$ are completely covered by Theorem III and Corollary II. We list below expressions for $\chi_r(\lambda)$ as a sum of irreducible characters $\{\mu\}$ for every other r, m, (λ) such that $M \leqq 10$.

$$M = 6,\ m = 3$$

$$\chi_2(3) = \{6\} + \{42\}$$

$$\chi_2(21) = \{42\} + \{321\} + \{2^3\} + \{31^3\}$$

$$\chi_2(1^3) = \{2^3\} + \{21^4\}$$

$$M = 8,\ m = 4$$

$$\chi_2(4) = \{8\} + \{62\} + \{4^2\}$$

$$\chi_2(31) = \{62\} + \{521\} + \{51^3\} + \{4^2\} + \{431\} + \{42^2\} + \{3^21^2\}$$

$$\chi_2(2^2) = \{4^2\} + \{42^2\} + \{3^21^2\} + \{2^4\}$$

$$\chi_2(21^2) = \{42^2\} + \{41^4\} + \{3^21^2\} + \{32^21\} + \{321^3\} + \{2^4\} + \{2^21^4\}$$

$$\chi_2(1^4) = \{2^4\} + \{2^21^4\} + \{1^8\}$$

$$M = 9,\ m = 3$$

$$\chi_3(3) = \{9\} + \{72\} + \{63\} + \{52^2\} + \{4^21\}$$

$$\begin{aligned}\chi_3(21) = {} & \{63\} + \{531\} + \{52^2\} + \{521^2\} + \{4^21\} + \{432\} + \{431^2\} \\ & + 2\{42^21\} + \{421^3\} + \{41^5\} + \{3^3\} + \{3^221\} + \{3^21^3\} + \{32^3\} \\ & + \{32^21^2\}\end{aligned}$$

$$\chi_3(1^3) = \{3^3\} + \{32^21^2\} + \{31^6\} + \{2^31^3\}$$

$$M = 10,\ m = 5$$

$$\chi_2(5) = \{10\} + \{82\} + \{64\}$$

$$\begin{aligned}\chi_2(41) = {} & \{82\} + \{721\} + \{71^3\} + \{64\} + \{631\} + \{62^2\} + \{541\} \\ & + \{531^2\} + \{4^22\}\end{aligned}$$

$$\begin{aligned}\chi_2(32) = {} & \{64\} + \{62^2\} + \{541\} + \{532\} + \{531^2\} + \{4^22\} + \{4321\} \\ & + \{42^3\} + \{3^31\}\end{aligned}$$

$$\begin{aligned}\chi_2(31^2) = {} & \{62^2\} + \{61^4\} + \{531^2\} + \{52^21\} + \{521^3\} + \{4^22\} + \{4321\} \\ & + \{431^3\} + \{42^3\} + \{421^4\} + \{3^221^2\}\end{aligned}$$

$$\begin{aligned}\chi_2(2^21) = {} & \{4^22\} + \{4321\} + \{431^3\} + \{42^3\} + \{3^31\} + \{3^221^2\} + \{32^31\} \\ & + \{32^21^3\} + \{2^5\}\end{aligned}$$

$$\begin{aligned}\chi_2(21^3) = {} & \{42^3\} + \{421^4\} + \{3^221^2\} + \{32^31\} + \{32^21^3\} + \{321^5\} + \{31^7\} \\ & + \{2^5\} + 2\{2^31^4\}\end{aligned}$$

$$\chi_2(1^5) = \{2^5\} + \{2^31^4\} + \{21^8\}.$$

CHAPTER III.

7. The Lie representation. Denote (as above) the partition of r into α_1 1's, α_2 2's, $\cdots$ by (α) and call $\alpha_0 = \alpha_1 + \alpha_2 + \cdots$ the *length* of the partition. We now arrange all partitions (α) of r in an arbitrary array $(\alpha^1), \cdots, (\alpha^{p(r)})$ subject only to the condition that (α) is to precede (β) if $\alpha_0 > \beta_0$. It is clear that any solution in non-negative integers of $r = i_1 + \cdots + i_{\alpha_0}$ defines a unique partition (α) of length α_0, and conversely that to each (α) there corresponds a unique partition $r = i_1 + \cdots + i_{\alpha_0}$ with $0 \leqq i_1 \leqq i_2 \leqq \cdots$.

Let $u_{i1}, \cdots, u_{i\psi_i}$ be a basis for $\boldsymbol{L}^i$, $i = 1, 2, \cdots$. E. Witt[24] has introduced parenthesis quantities $(u_1, \cdots, u_k)$, $k = 1, 2, \cdots$ with elements u_j in $\boldsymbol{L}$ with the following properties:

$$(7.1) \qquad (\cdots, u, v, \cdots) - (\cdots, v, u, \cdots) = (\cdots, uv, \cdots)$$

$$(7.2) \qquad (\cdots, \Sigma a_j u_j, \cdots) = \Sigma a_j (\cdots, u_j, \cdots)$$

$$(7.3) \qquad (\cdots, u)(v, \cdots) = (\cdots, u, v, \cdots)$$

He then proved that

a) Under the new multiplication defined by 7.3 the set of parenthesis quantities constitute a ring $\boldsymbol{R}$ simply isomorphic with the free non-commutative ring $\boldsymbol{X}$ under the correspondence $x_i \rightarrow (s_i)$.

b) The $\boldsymbol{K}$-module $\boldsymbol{R}^r$ has as a basis the set of all *ordered* parenthesis quantities $(u_{i_1 j_1}, \cdots, u_{i_m j_m})$ where (i) $i_1 \leqq i_2 \leqq \cdots$, (ii) $j_k \leqq j_{k+1}$ if $i_k = i_{k+1}$, (iii) $i_1 + \cdots + i_m = r$.

If U is any parenthesis quantity which is made up solely of basis elements u_{ij} we associate with it the ordered parenthesis element U^* made up from the same elements u_{ij} arranged so as to satisfy bi) and bii). E. g. if $U = (u_{4,1}, u_{3,5}, u_{3,2}, u_{5,1})$ then $U^* = (u_{3,2}, u_{3,5}, u_{4,1}, u_{5,1})$. If $U = (u_{i_1 j_1}, \cdots, u_{i_m j_m})$ we say that U *belongs to the partition* (α) defined by $i_1 + \cdots + i_m = r$. By repeated applications of 7.1 and 7.2 we can show that $U - U^* = U^{**}$ is a linear combination of parenthesis quantities all of degree r and belonging to partitions of length less than $\alpha_0 = m$.

Returning now to the matrices $L_r(A)$, we arrange the basis elements of $\boldsymbol{R}^r$ into $p(r)$ sets, a set consisting of all ordered parenthesis quantities belonging to a given partition (α), and arrange these according to the ordering of the partitions to which they belong. The automorphism A: $(s_i) \rightarrow (s'_i) = \Sigma a_{ik}(s_k)$ replaces the ordered parenthesis quantity U by $U' = (u'_{i_1 j_1}, \cdots, u'_{i_m j_m})$.

---

[24] [7] pp. 153-155.

Denote by $\Pi_r(A)^*$ the matrix whose rows are the coefficients in the expressions for the U' in terms of the ordered parenthesis quantities U. Since $\boldsymbol{R}$ is isomorphic to $\boldsymbol{X}$ the representation $A \to \Pi_r(A)^*$ is equivalent to the Kronecker r-th power representation $A \to \Pi_r(A)$ and so its irreducible constituents are known.[25]

Set $[i] = \operatorname{tr} L_i(A)$. Denote by $A[\alpha]$ the matrix

$$S_{\alpha_1}(L_1(A)) \times S_{\alpha_2}(L_2(A)) \times \cdots$$

and set $[(\alpha)] = \operatorname{tr} A[\alpha]$. Our goal is a formula which expresses $\operatorname{tr} \Pi_r(A)$ in terms of the characters $[(\alpha)]$ of $GL(n)$. Such a formula can then be regarded as a recursion formula from which one can compute $[r]$ as a function of the $[(\alpha)]$ (with $\alpha_0 > 1$) and $\operatorname{tr} \Pi_r(A)$.

The arrangement of the basis elements of $\boldsymbol{R}^r$ into $p(r)$ sets induces a refinement of $\Pi_r(A)^*$ into $p(r)^2$ submatrices $A[(\alpha^i)(\alpha^j)]$, $i, j = 1, \cdots, p(r)$. The following theorem shows that $\Pi_r(A)^*$ has zeros below blocks along the main diagonal and thus enables us to compute traces using only these diagonal blocks.

THEOREM VI. (i) $A[(\alpha^i)(\alpha^j)] = 0$ *if* $i > j$

(ii) $A[(\alpha^i)(\alpha^i)] = A[\alpha^i]$.

Proof. Let $U = (u_{i_1 j_1}, \cdots, u_{i_m j_m}) = U(j_1, \cdots, j_m)$ belong to the partition (α). Then substituting $u'_{ij} = \Sigma a^{(i)}_{j\rho} u_{i\rho}$ in $U' = (u'_{i_1 j_m}, \cdots, u'_{i_m j_m})$ and applying 7.2 we express U' in the form

$$U' = \sum_{(\rho)} a^{(i_1)}_{j_1\rho_1} \cdots a^{(i_m)}_{j_m\rho_m}\; U(\rho_1, \cdots, \rho_m). \tag{7.4}$$

Observe that each $U(\rho_1, \cdots, \rho_m)$ belongs to the partition (α). We have already established an expression $U(\rho_1, \cdots, \rho_m) = U^*(\rho_1, \cdots, \rho_m) + U^{**}(\rho_1, \cdots, \rho_m)$ where U^{**} is a linear combination of ordered parenthesis elements belonging to partitions of length less than $m = \alpha_0$, and where U^* is an ordered parenthesis element belonging to the partition (α). This establishes part (i) of the theorem.

Before treating (ii) in general we consider two special cases. Suppose that $i_1 < i_2 < \cdots$. Then each $U(\rho_1, \cdots, \rho_m)$ is already an ordered parenthesis element and $A[(\alpha^i)(\alpha^i)] = \| a^{(\alpha^i)}(j_1, \cdots, j_m;\ \rho_1, \cdots, \rho_m) \|$ has $a^{(\alpha^i)}(j_1, \cdots, j_m; \rho_1, \cdots, \rho_m) = a^{(i_1)}_{j_1\rho_1} \cdots a^{(i_m)}_{j_m\rho_m}$, which is one way of defining the Kronecker product $L_{i_1}(A) \times \cdots \times L_{i_1}(A)$.

The second special case is $i_1 = i_2 = \cdots = I$. Then $U^*(\rho_1, \cdots, \rho_m) = U(k_1, \cdots, k_m)$ where $k_1, \cdots, k_m$ are the integers $\rho_1, \cdots, \rho_m$ written in increasing (natural) order. Hence in this case $a^{(\alpha')}(j_1,\cdots,j_m;k_1,\cdots,k_m) = \Sigma a^{(I)}_{j_1\rho_1}\cdots\cdots a^{(I)}_{j_m\rho_m}$ where the sum extends over all distinct rearrangements $\rho_1,\cdots,\rho_m$ of $k_1,\cdots,k_m$. But this is just the definition of $S_m L_I(A)$.

---

[25] [3] p. 119 or [4] p. 213.

The situation for an arbitrary partition (α) is now readily handled by combining these two special cases. We apply the arguments of the second case wherever any $\alpha_j > 1$ and then apply the arguments of the first case to the parts thus obtained to get the theorem as claimed.

Since $\operatorname{tr}\Pi_r(A)^* = \operatorname{tr}\Pi_r(A) = \{1\}^r$ we obtain upon adding the traces of all the matrices $A[\alpha]$ the formula

THEOREM VII. $\{1\}^r = \Sigma[(\alpha)]$ *the summation being over all partitions* $\alpha_1 + 2\alpha_2 + \cdots + r\alpha_r = r$ *of* r.

8. Simplification of the recursion formula and the analyses of the Lie representations $L_r(A)$ with $r \leqq 10$. For the purposes of this section it is convenient to use another notation for partitions. The partition labelled (α) in the preceding section will now be denoted by $(\lambda) = (\lambda_1,\cdots,\lambda_k)$ where $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_k \geqq 0$, $k = \alpha_0$ and α_i is the number of λ's equal to i. We shall designate $[(\alpha)]$ by $[\lambda]$. We shall use the following properties of the functions $[\lambda]$.

(8.1) $\{1\}^s = \Sigma[\lambda]$ (summation over all partitions (λ) of s.)

(8.2) $[1^s] = \operatorname{tr} S_s(L_1(A)) = \{s\}$.

(8.3) $\lambda_v > \lambda_{v+1}$ implies $[\lambda] = [\lambda_1,\cdots,\lambda_v][\lambda_{v+1},\cdots,\lambda_k]$.

A non-increasing sequence of positive integers, $\lambda_1,\cdots,\lambda_v$ will be called *r-proper* if $1 < \lambda_v < r - (\lambda_1 + \cdots + \lambda_v)$. Let (λ) be a partition of s. If $s < r$, (λ) will be called *r-proper* if each of the sequences $\lambda_1,\cdots,\lambda_v$ for $v = 1,\cdots,k$ is *r-proper*. If $s = r$, (λ) will be called *r-proper* if $k > 1$ and the sequences $\lambda_1,\cdots,\lambda_v$ for $v = 1,\cdots,k-1$ are all *r-proper*. All other partitions are said to be *r-improper*.

THEOREM VIII. $[r] = \sum_{p=-1}^{r-1}\{p,1\}f_r(s)$ *where* $f_r(s) = \Sigma[\lambda]$ *the sum extending over all r-proper partitions* (λ) *of* $s = r - p - 1$.

Proof. We solve for $[r]$ in the equation obtained by equating the right hand side of 8.1 for $s = r$ to $\{1\}$ times the right hand side of 8.1 for $s = r - 1$. The first step is a division of the partitions of r into classes. The first class C_1 consists of all r-proper partitions. The second class C_2 consists of all r-improper partitions having $k > 0$ and some $\lambda_v > r - (\lambda_1 + \cdots + \lambda_v)$. The third class C_3 contains all the remaining r-improper partitions of r save

$r = r$, which is the sole element of the fourth class $\boldsymbol{C}_4$. The partitions of $r - 1$, all necessarily r-improper, are divided into two classes. $\boldsymbol{C}'_2$ consists of all partitions of $r - 1$ which have some $\lambda_v > r - (\lambda_1 + \cdots + \lambda_v)$ and $\boldsymbol{C}'_3$ contains all other partitions of $r - 1$.

The partitions in $\boldsymbol{C}_3$ and $\boldsymbol{C}'_3$ can be put into one to one correspondence by ordering to each partition in $\boldsymbol{C}'_3$ the partition in $\boldsymbol{C}_3$ obtained by appending a single 1. Denote by $\boldsymbol{C}_3(p)$, $\boldsymbol{C}'_3(p)$ the set of partitions in $\boldsymbol{C}_3$, $\boldsymbol{C}'_3$ which end in $p + 1$, p ones, respectively.

LEMMA VIII. *Denote by $G_3(p)$, $G'_3(p)$ the sum of all $[\lambda]$ for which (λ) lies in $\boldsymbol{C}_3(p)$, $\boldsymbol{C}'_3(p)$, respectively. Then $G'_3(p)\{1\} - G_3(p) = \{p, 1\}f_r(s)$ where $s = r - p - 1$, and $p > 0$.*

Proof. Let $(\lambda) = (\lambda_1, \cdots, \lambda_{k-p-1}, 1, \cdots, 1)$ lie in $\boldsymbol{C}_3(p)$ and let (λ') be (λ) with the last 1 deleted. Then by 8.3 and 8.2 $[\lambda] = [\lambda_1, \cdots, \lambda_{k-p-1}]\{p + 1\}$ and $[\lambda'] = [\lambda_1, \cdots, \lambda_{k-p-1}]\{p\}$. Therefore $[\lambda']\{1\} - [\lambda] = [\lambda'']\{p, 1\}$; where, since (λ) is not in $\boldsymbol{C}_2$, $\lambda_{k-p-1} > 1$ implies that $(\lambda'') = (\lambda_1, \cdots, \lambda_{k-p-1})$ is an r-proper partition of $s = r - p - 1$. Furthermore, it being clear that each r-proper partition (λ'') of s will occur exactly once as a factor of $\{p, 1\}$ in the difference $G'_3(p)\{1\} - G_3(p)$, this establishes our lemma.

LEMMA IX. *Denote by G_2, G'_2 the sum of all $[\lambda]$ for which (λ) lies in $\boldsymbol{C}_2$, $\boldsymbol{C}'_2$ respectively. Then $G'_2\{1\} - G_2 = 0$.*

Proof. Let (μ) belong to $\boldsymbol{C}_2$ and suppose that v is the smallest integer for which $\mu_v > s = r - (\mu_1 + \cdots + \mu_v)$. Then let $H(\mu)$ be the sum of all $[\lambda]$ for which $(\lambda) = (\mu_1, \cdots, \mu_v, \lambda_{v+1}, \cdots)$ is in $\boldsymbol{C}_2$. Then since $\mu_v > \lambda_{v+1} + \cdots \geqq \lambda_{v+1}$ we have by 8.3 that $[\lambda] = [\mu_1, \cdots \mu_v][\lambda_{v+1}, \cdots]$. Hence, by 8.1, $H(\mu) = [\mu_1, \cdots, \mu_v]\{1\}^s$. If $H'(\mu)$ denotes the sum of all $[\lambda']$ for which $(\lambda') = (\mu_1, \cdots, \mu_v, \lambda'_{v+1}, \cdots)$ is in $\boldsymbol{C}'_2$ we have similarly $H'(\mu) = [\mu_1, \cdots, \mu_v]\{1\}^{s-1}$. Now evidently G_2 can be written as a sum of certain $H(\mu)$ and G'_2 is then the sum of the corresponding $H'(\mu)$. From this the lemma follows.

Let G_1 be the sum of all $[\lambda]$ for (λ) in $\boldsymbol{C}_1$. Evidently $G_1 = f_r(r)$. The theorem now follows from the equality

$$(G'_2 + G'_3)\{1\} = [r] + G_1 + G_2 + G_3$$

and the relations $\{0, 1\} = 0$, $\{-1, 1\} = -\{0\} = -1$.

We now list the analyses of $[r]$ for $r \leqq 10$.

$[1] = \{1\}$; $[2] = \{1^2\}$; $[3] = \{21\}$; $[4] = \{31\} + \{21^2\}$;

$[5] = \{41\} + \{32\} + \{31^2\} + \{2^21\} + \{21^3\}$

$$[6] = \{51\} + \{42\} + 2\{41^2\} + \{3^2\} + 3\{321\} + \{31^3\} + 2\{2^21^2\} + \{21^4\}$$

$$\begin{aligned}[7] = {}& \{61\} + 2\{52\} + 2\{51^2\} + 2\{43\} + 5\{421\} + 3\{41^3\} + 3\{3^21\} \\ &+ 3\{32^2\} + 5\{321^2\} + 2\{31^4\} + 2\{2^31\} + 2\{2^21^3\} + \{21^5\}\end{aligned}$$

$$\begin{aligned}[8] = {}& \{71\} + 2\{62\} + 3\{61^2\} + 4\{53\} + 8\{521\} + 4\{51^3\} + \{4^2\} \\ &+ 9\{431\} + 6\{42^2\} + 12\{421^2\} + 4\{41^4\} + 6\{3^22\} + 6\{3^21^2\} \\ &+ 9\{32^21\} + 8\{321^3\} + 3\{31^5\} + \{2^4\} + 4\{2^31^2\} + 2\{2^21^4\} + \{21^6\}\end{aligned}$$

$$\begin{aligned}[9] = {}& \{81\} + 3\{72\} + 3\{71^2\} + 5\{63\} + 12\{621\} + 6\{61^3\} + 5\{54\} \\ &+ 18\{531\} + 13\{52^2\} + 21\{521^2\} + 8\{51^4\} + 9\{4^21\} + 19\{432\} \\ &+ 24\{431^2\} + 24\{42^21\} + 21\{421^3\} + 6\{41^5\} + 4\{3^3\} + 19\{3^221\} \\ &+ 13\{3^21^3\} + 9\{32^3\} + 18\{32^21^2\} + 12\{321^4\} + 3\{31^6\} + 5\{2^41\} + \\ &+ 3\{2^21^5\} + \{21^7\}\end{aligned}$$

$$\begin{aligned}[10] = {}& \{91\} + 3\{82\} + 4\{81^2\} + 8\{73\} + 15\{721\} + 7\{71^3\} + 8\{64\} \\ &+ 31\{631\} + 20\{62^2\} + 34\{621^2\} + 11\{61^4\} + 5\{5^2\} + 29\{541\} \\ &+ 45\{532\} + 53\{531^2\} + 51\{52^21\} + 43\{521^3\} + 13\{51^5\} + 22\{4^22\} \\ &+ 31\{4^21^2\} + 21\{43^2\} + 73\{4321\} + 49\{431^3\} + 26\{42^3\} + \{42^21^2\} \\ &+ 31\{421^4\} + 8\{41^6\} + 18\{3^31\} + 25\{3^22^2\} + 39\{3^221^2\} + 21\{3^21^4\} \\ &+ 25\{32^31\} + 26\{32^21^3\} + 13\{321^5\} + 2\{31^7\} + 2\{2^5\} + 8\{2^41^2\} \\ &+ 5\{2^31^4\} + 3\{2^21^6\} + \{21^8\}.\end{aligned}$$

INSTITUTE FOR ADVANCED STUDY.

BIBLIOGRAPHY.

[1] D. E. Littlewood, *The Theory of Group Characters and Matrix Representations of Groups*, Oxford (1940).

[2] P. A. MacMahon, *An Introduction to Combinatory Analysis*, Cambridge (1920).

[3] F. D. Murnaghan, *The Theory of Group Representations*, Baltimore (1938).

[4] H. Weyl, *The Classical Groups*, Princeton (1939).

[5] G. deB. Robinson, "On the representations of the symmetric group," *American Journal of Mathematics*, vol. 60 (1938), pp. 745-760.

[6] J. H. M. Wedderburn, "The functional equation . . . ," *Annals of Mathematics*, vol. 24 (1922-23), pp. 121-140.

[7] E. Witt, "Treue Darstellung Liescher Ringe," *Journal fur die Reine und angewandte Mathematik*, vol. 177 (1937), pp. 152-160.

COMPOSITE POLYNOMIALS WITH COEFFICIENTS IN AN ARBITRARY FIELD OF CHARACTERISTIC ZERO.*

By HOWARD LEVI.

A polynomial $F(x)$ is said to be *composite* if there exist polynomials $P(y)$ and $Q(x)$, each of whose degrees exceeds unity, and such that

$$F(x) = P(Q(x)).$$

Under these circumstances PQ is called a *decomposition* of F. If no such pair exists F is said to be *prime*.

Composite polynomials whose coefficients are complex numbers were investigated by J. F. Ritt.[1] His results completely describe the decompositions such polynomials may possess. H. T. Engstrom[2] discussed the case in which the coefficient domain of the polynomials is any field of characteristic zero. He proved that the first of Ritt's two chief results is valid for polynomials whose coefficients are elements of such a field. The present paper is devoted to showing that Ritt's second result is likewise valid for the abstract case.

Ritt's main conclusions may be described as follows. He first proved that if a polynomial has two decompositions into prime polynomials (obviously it always has at least one); that is if $F = P_1P_2 \cdots P_r$ and $F = Q_1Q_2 \cdots Q_s$, then $r = s$, and the degrees of the polynomials P_i are the same (as a set with no reference to order) as those of the Q_j. Thus the situation bears a certain resemblance to Jordan-Hölder-like phenomena met in algebra. However there is also a striking difference. Unlike the situation in such a typical instance as that of composition series in groups, where the constituent normal subgroups can be extremely variegated, in decompositions of polynomials there is practically no latitude. Ritt's second result is that if $F = P_1P_2 = Q_1Q_2$ where the P_i and Q_i are prime, then, if P_2 and Q_2 are not merely linear transforms of each other, there are essentially only two possibilities for F. They are:

I. $$F(x) = (x^n)^r[g(x^n)]^n = [x^r g(x^n)]^n,$$

where clearly

$$P_1(y) = y^n, \qquad P_2(x) = x^r g(x^n),$$

---

* Received May 29, 1941.

[1] J. F. Ritt, "Prime and composite polynomials," *Transactions of the American Mathematical Society*, vol. 23 (1922), pp. 51-66.

[2] H. T. Engstrom, "Polynomial Substitutions," *American Journal of Mathematics*, vol. 63 (1941), pp. 249-255.

and

$$Q_1(y) = y^r g(y), \quad Q_2(x) = x^n.$$

This includes as a special case

$$F(x) = (x^r)^n = (x^n)^r.$$

II. Let $T_r(x)$ be defined by $T_r(\cos A) = \cos rA$.
Then

$$F(x) = T_m(T_n(x)) = T_n(T_m(x)).$$

These exhaust the possibilities in the sense that any F of the type discussed above which is not included in I or II becomes included when each of x, P_2, Q_2, F is replaced by a suitable linear integral transform of itself. It was further pointed out how these two polynomial results give a complete description of the general case. We shall return to this point later in the paper.

The first result listed above has been shown by Engstrom to hold for an arbitrary coefficient field of characteristic zero. In this paper it is shown that the second also holds for such fields.

Statement of the Problem.

1. Let K be any field of characteristic zero and let x be an indeterminate. A polynomial $F(x)$ with coefficients in K will be called *composite in* K if F admits a decomposition PQ as above where P and Q are polynomials whose coefficients are elements of K.

2. It is readily seen that if F is composite in any extension K^* of K over which x is an indeterminate, then F is composite in K. To establish this let $F = PQ$ where P and Q have coefficients in K^*. Let $Q^* = aQ + b$ where a and b are elements of K^* chosen so that the leading coefficient of Q^* is unity and its constant term zero. Let $F^* = cF$ where c is an element of K chosen so that the leading coefficient of F^* is unity. Clearly $F^* = P^*Q^*$ where P^* is a polynomial. Then if

$$Q^* = x^n + \sum_{i=1}^{n-1} a_i x^{n-i}; \qquad F^* = x^{mn} + \sum_{j=1}^{mn} b_j x^{mn-j},$$

we have [3]

$$b_j = G_j(a_1, \cdots, a_{j-1}) + m a_j, \qquad j = 1, \cdots, n-1, \tag{1}$$

where G_j is an isobaric polynomial in $a_1, \cdots, a_{j-1}$ of weight j whose coefficients are integers. Thus a_i; $i = 1, \cdots, n-1$ is in K, so that F^* and therefore F is composite in K.

---

[3] Cf., Engstrom, *loc. cit.*

From equations (1) we see likewise that if $F = P_1P_2 = Q_1Q_2$ and if P_2 and Q_2 have the same degree, they must be linear functions of each other. This distinguishes the decomposition of polynomials from that of fractional rational functions.

3. We shall denote the field obtained by adjoining $a, b, \cdots, f$, to K by the expression $K(a, b, \cdots, f)$. Our problem may be treated from the point of view that $F(x) = P(Q(x))$ implies $K(F) \subset K(Q) \subset K(x)$. Conversely, it is known that if $K(F) \subset \Sigma \subset K(x)$, where Σ is a field, then Σ is generated by a polynomial with coefficients in K and F is composite in K.[4] Thus every decomposition $P_1P_2 \cdots P_r$ of F determines a chain of fields

$$K(P_1P_2 \cdots P_r) \subset K(P_2 \cdots P_r) \subset \cdots \subset K(P_r) \subset K(x)$$

all generated by polynomials, and conversely. Each field $K(P_i \cdots P_r)$ is an algebraic extension of its predecessor $K(P_{i-1}P_i \cdots P_r)$. It may be obtained by adjoining to the latter a root of the equation for u,

$$P_{i-1}(u) - P_{i-1}P_i \cdots P_r = 0.$$

The coefficients of this equation are in $K(P_{i-1}P_i \cdots P_r)$ and it is irreducible in this field.[5] The degree of this equation and of the extension is obviously that of P_{i-1}.

4. Let F have two decompositions into prime polynomials. Each determines a chain of fields, as described above. We have

$$K(F) \subset \Sigma_1 \subset \cdots \subset \Sigma_i \subset \cdots \subset K(x)$$

and

$$K(F) \subset \Sigma'_1 \subset \cdots \subset \Sigma'_i \subset \cdots \subset K(x)$$

where in each chain each field is maximally incident in its successor. If $\Sigma_i = \Sigma'_i$ for each i the chains are indentical and the two decompositions are essentially the same.[6] If this is not the case let i be the smallest integer for which $\Sigma_i \neq \Sigma'_i$. Understanding that $\Sigma_0 = \Sigma'_0 = K(F)$ we have $\Sigma_{i-1} = \Sigma'_{i-1} = \Sigma_i \wedge \Sigma'_i$. Both Σ_i and Σ'_i are algebraic extensions of Σ_{i-1}. Let their respective degrees over Σ_{i-1} be m and n. Engstrom has shown that these integers must be relatively prime.[7] Let $\Sigma \subset K(x)$ be the smallest field which contains Σ_i and Σ'_i. The degree of Σ over Σ_{i-1} is readily seen to be mn. In fact its degree over Σ_{i-1} must be a multiple of both m and n and consequently of mn. On the other hand Σ may be obtained by adjoining the generator of Σ'_i to Σ_i, which may be accomplished by solving an equation of degree n with coefficients in $\Sigma_{i-1} \subset \Sigma_i$. Thus the degree of Σ over Σ_{i-1} does not exceed mn and so must be exactly mn.

---

[4] Lueroth's theorem asserts the existence of a single generator for Σ which is, in general, a rational fractional function of x. Engstrom (*loc. cit.*) has shown that here a polynomial generator exists.

[5] Because it is of the first degree in the generator of the field. Cf. Van der Waerden, *Moderne Algebra*, vol. I, p. 207.

[6] One can secure new decompositions from any one by simply subjecting the arguments of the constituent polynomials to linear transformations. This process leaves the associated chain of fields unaltered.

It follows quickly that there is no field between Σ'_i and Σ, or between Σ_i and Σ. The choice of the integer i insures that there is no field between Σ_i and Σ_{i-1}, or between Σ'_i and Σ_{i-1}. Then by Engstrom's result, in the chain $\Sigma_{i-1} \subset \Sigma'_i \subset \cdots \subset \Sigma$ one of the fields has to be of degree m over its predecessor, since this is true of the chain $\Sigma_{i-1} \subset \Sigma_i \subset \cdots \subset \Sigma$. Clearly the existence of a field between Σ and Σ'_i would require the degree of Σ over Σ_{i-1} to exceed mn. Similarly there is no field between Σ and Σ_i.

5. We are led to the following two polynomial problem. *Given*

$$F(x) = P(p(x)) = Q(q(x))$$

where F, P, p. Q, q are polynomials with coefficients in K, P and q are prime and of degree n, p and Q are prime and of degree m. F is composite and of degree mn. The integers m and n are relatively prime and each exceeds unity. It is required to determine the possibilities for F, P, p, Q, q.

From now on the only use we make of the fact that m and n are relatively prime is that they are then unequal. We suppose $m > n$.

The field situation may be represented as follows.

$$K(F) = K(p) \cap K(q) \begin{array}{c} \subset K(p) \subset \\ \subset K(q) \subset \end{array} K(p, q) = K(x),$$

where the incidence is maximal in each case. We may suppose that K contains the n-th roots of unity and the roots of $F'(u) = 0$ without disturbing these relations. We may further suppose that the leading coefficient of each of the five polynomials is unity, and that the constant term of each is zero.

Relations governing the Factorization of $F'(x)$.

6. It will be useful to list the irreducible equations satisfied by the generators of the various fields. We begin with x which satisfies $F(u) - F(x) = 0$, $p(u) - p(x) = 0$, $q(u) - q(x) = 0$, irreducible respectively in $K(F)$, $K(p)$, $K(q)$. The irreducible equation with coefficients in $K(F)$ satisfied by p is $P(u) - F(x) = 0$. Since $F = Qq$, the coefficients of $P(u) - F(x)$

---

[7] Engstrom, *loc. cit.*

$= 0$ are contained in $K(q)$. We need the irreducibility of this equation in $K(q)$. This follows from the fact that the adjunction of a root of this equation to $K(q)$ yields the same field as that obtained by adjoining a root of $q(u) - q(x) = 0$ to $K(q)$, namely $K(x)$, and this last equation, which we know to be irreducible in $K(q)$, has the same degree as $P(u) - F(x) = 0$. Similarly, q satisfies $Q(u) - F(x) = 0$, which is irreducible in $K(p)$ as well as in $K(F)$.

7. Let c be any element of K. Because the leading coefficient of F is unity, the norm[8] of $x - c$ relative to $K(F)$ is $(-1)^{mn+1}(F(x) - F(c))$. We write

$$\mathfrak{N}_{x|F}(x - c) = F(x) - F(c),$$

using the notation $\mathfrak{N}_{x|F}(w)$ to designate either the norm of w relative to $K(F)$ or its negative, it being understood that w is an element of $K(x)$. (This seeming ambiguity should lead to no confusion, since, in every case in which we employ the notation, multiplication of $\mathfrak{N}_{x|F}(w)$ by a known power of (-1) converts it into the true norm. Which sign is used is irrelevant to our argument). Similarly, with obvious interpretation for the symbols $\mathfrak{N}_{x|p}(w)$, etc.,

$$\mathfrak{N}_{x|p}(x - c) = p(x) - p(c), \qquad \mathfrak{N}_{x|q}(x - c) = q(x) - q(c),$$
$$\mathfrak{N}_{p|F}(p - c) = P(p) - P(c) = F(x) - P(c),$$
$$\mathfrak{N}_{q|F}(q - c) = Q(q) - Q(c) = F(x) - Q(c).$$

In addition, making use of the irreducibility of $P(u) - F(x) = 0$ in $K(q)$ we have

$$\mathfrak{N}_{x|q}(p - c) = \mathfrak{N}_{p|F}(p - c) = F(x) - P(c),$$

and similarly

$$\mathfrak{N}_{x|p}(q - c) = \mathfrak{N}_{q|F}(q - c) = F(x) - Q(c).$$

Finally we observe that if w is in $K(p)$, $\mathfrak{N}_{x|p}(w) = w^m$ and if w is in $K(q)$, $\mathfrak{N}_{x|q}(w) = w^n$.

8. $F'(x)$ is in $K(x)$ and in that field splits up into the product of $mn - 1$ linear factors. We have $F'(x) = P'(p)p'(x) = Q'(q)q'(x)$. Thus $\mathfrak{N}_{x|p}(F'(x))$ may be computed in two ways. We have

$$\mathfrak{N}_{x|p}(F(x)) = \mathfrak{N}_{x|p}(P'(p) \ (P'(p)) \ \mathfrak{N}_{x|p}(p'(x)) = P'(p)^m \mathfrak{N}_{x|p}(p'(x) \ (p'(x))$$

and

$$\mathfrak{N}_{x|p}(F'(x)) = \mathfrak{N}_{x|p}(Q'(q) \ (Q'(q)) \ \mathfrak{N}_{x|p}(q'(x)) = \mathfrak{N}_{q|F}(Q'(q)) \ \mathfrak{N}_{x|p}(q'(x)).$$

---

[8] As defined in Van der Waerden, *loc. cit.*, p. 135. Note that the norm of a product is the product of the norms, and that a polynomial divides its norm.

The equation

$$(2)\qquad P'(p)^m\mathfrak{N}_{x|p}(p'(x)) = \mathfrak{N}_{q|F}(Q'(q)\ (Q'(q))\ \mathfrak{N}_{x|p}(q'(x))$$

and the analogous one obtained by evaluating $\mathfrak{N}_{x|p}(F'(x))$ are fundamental in our work. A study of the factorization of the different members of these equations will yield a solution of the two polynomial problem. While it is perhaps not immediately obvious, this procedure is closely related to Ritt's work. The evalution of the various norms is similar in effect to the excursions around the critical points of the inverse functions which figured so prominently in his paper.

9. We single out any factor $p - a$ of $P'(p)$ and examine the way it figures in (2). Its multiplicity in the left member is $wm + r$, where w is the multiplicity of $p - a$ in $P'(p)$ and where r is the number of linear (in x) factors of $p'(x)$ which divide $p - a$. To find its multiplicity in the right member, observe that $\mathfrak{N}_{x|p}(q'(x))$ contributes t, where t is the number of linear (in x) factors of $q'(x)$ which divide $p - a$. Let $p - a$ divide $F - A$. Its multiplicity in $F - A$ must be $w + 1$. If there are s linear (in q) factors of $Q'(q)$ which divide $F - A$, then the multiplicity of $p - a$ in $\mathfrak{N}_{q|F}(Q'(q))$ is $s(w + 1)$. Collecting our results, we have

$$(3)\qquad mw + r = s(w + 1) + t.$$

Since $m > s$ (the degree of Q' in q is $m - 1$) it follows that

$$(4)\qquad m \leq 2s + t.$$

We are going to use this inequality to show that all the factors $p - a$ of $P'(p)$ divide at most two distinct polynomials $F - A$ and $F - A'$. Suppose this were not the case, and let $p - a$, $p - a'$, $p - a''$ be factors of $P'(p)$ which divide $F - A$, $F - A'$, $F - A''$ respectively, where A, A', A'' are three distinct elements of K. We have $m \leq 2s' + t'$ and $m \leq 2s'' + t''$ where the inequalities and the primed symbols are obtained as above. By adding the inequalities we obtain

$$3m \leq 2(s + s' + s'') + t + t' + t''.$$

Now $s + s' + s''$ does not exceed the degree of $Q'(q)$ in q, namely $m - 1$ and $t + t' + t''$ does not exceed the degree of $q'(x)$ in x, namely $n - 1 < m - 1$. This inequality implies the absurd result that $3m < 3(m - 1)$.

Derivation of Case I.

10. Let us first treat the case in which all the factors of $P'(p)$ are contained in a single $F - A$. We must have

$$F-A=(p-a_1)^{w_1+1}\cdots(p-a_r)^{w_r+1}$$

where the w_i are non-negative integers with $\Sigma w_i=n-1$, $\Sigma(w_i+1)=n$. Clearly $r=1$, $w_1=n-1$ whence

$$F-A=(p-a_1)^n. \tag{5}$$

In order for $P(p)$ to be prime the integer n must be prime. We now show how this result determines $q(x)$ as a linear transform of x^n. It will then be easy to show that F comes under I.

11. The field $K(x)$ may be obtained from $K(q)$ by solving the irreducible (in $K(q)$) equation $P(u)-F(x)=0$. Thus, in view of (5), $K(x)$ is a cyclic extension of $K(q)$. We show that the n roots of the equation $q(u)-q(x)=0$ must be linear integral functions of x. They must be rational functions of x since $K(x)$ is a normal extension of $K(q)$; they must be linear fractional functions of x because they also generate $K(x)$; and finally they must be entire functions of x because otherwise they could not satisfy $q(u)-q(x)=0$. Let the $n-1$ roots of this equation other than x be

$$x_i=d_ix+e_i, \qquad d_i, e_i\in K, \qquad i=2,\cdots,n.$$

The "Lagrange Resolvent"[9]

$$R=x+\sum_{i=2}^{n}\omega^{i-1}x_i$$

(ω is a primitive n-th root of unity) therefore is of the form $Dx+E$ where D and E are elements of K. Since R does not belong to $K(q)$ we know that D is not zero. It is a characteristic property of R that $R^n\in K(q)$ and our knowledge about the degree of R and of $q(x)$ permits us to write

$$R^n=gq(x)+h, \qquad g,h\in K.$$

By subjecting the generators of the fields to appropriate linear integral transformations, we may suppose that (old notation retained)

$$F(x)=(p(x))^n, \qquad q(x)=x^n,$$

where $p(x)$ and $F(x)$ have unity as leading coefficients and where the constant term of $p(x)$ is zero. We then confront the problem of determining $Q(y)$ and $p(x)$ in such a fashion that $F(x)=(p(x))^n=Q(x^n)$. Suppose that

$$p(x)=x^m+\sum_{i=1}^{m-1}a_ix^{m-i},$$

$$F(x)=x^{mn}+\sum_{i=1}^{m-1}b_ix^{mn-i}.$$

---

[9] Van der Waerden, *loc. cit.*, p. 177.

Equations (1) show that a_i must be zero for $i \not\equiv 0(n)$. Then

$$p(x) = x^r g(x^n)$$

where r is an integer less than n and g is a polynomial with coefficients in K. Since $p(x)$ is prime r must be different from zero. Clearly

$$Q(y) = y^r g(y)^n, \qquad F(x) = (x^r g(x^n))^n.$$

Derivation of Case II.

12. We turn now to the case in which $P'(p)$ has non-constant factors in common with $F - A$ and $F - A'$, where A and A' are distinct elements of K. We have

$$\mathfrak{N}_{p|F}(P'(p)) = c(F - A)^v (F - A')^{v'}$$

with $v + v' = n - 1$ and $vv' \neq 0$. The symbol c denotes some element of K. The discussion is divided into two parts. We consider the equation

$$(6) \quad \mathfrak{N}_{x|q}(F'(x)) = Q'(q)^n \mathfrak{N}_{x|q}(q'(x)) = c(F - A)^v (F - A')^{v'} \mathfrak{N}_{x|q}(p'(x))$$

and assume first that $F - A$ and $F - A'$ have factors $q - b$ and $q - b'$ respectively which are prime to $Q'(q)$. We proceed to show that this leads to the case of the trigonometric polynomials described under II. Equation (6) reveals that our assumption about $q - b$ and $q - b'$ implies that $(q - b)^v (q - b')^{v'}$ divides $\mathfrak{N}_{x|q}(q'(x))$. Because the degree of $\mathfrak{N}_{x|q}(q'(x))$ in q is $n - 1 = v + v'$, there is precisely one such $q - b$ and one such $q - b'$. Let us see what is involved in the fact that $F - A$ has only one factor $q - b$ prime to $Q'(q)$. From

$$F - A = (q - b)[(q - b_1)^{w_1+1} \cdots (q - b_r)^{w_r+1}],$$

with each $w_i > 0$, we have $\Sigma(w_i + 1) = m - 1$, whence $\Sigma w_i \geq (m - 1)/2$. That is, the h. c. f. of $Q'(q)$ and $F - A$ has a degree in q not less than $(m - 1)/2$. Similarly, the h. c. f. of $Q'(q)$ and $F - A'$ is of degree not less than $(m - 1)/2$ in q. Since the total degree of $Q'(q)$ in q is $m - 1$, each of these degrees must be exactly $(m - 1)/2$. This implies that all the w_i (and the w'_i which figure in $F - A'$) must be unity. Therefore

$$m^2(F - A)(F - A') = (q - b)(q - b')Q'(q)^2.$$

The linear transformations

$$F = \tfrac{1}{2}(A' - A)F^* + \tfrac{1}{2}(A' + A)$$
$$q = \tfrac{1}{2}(b' - b)q^* + \tfrac{1}{2}(b' + b)$$

are permissible, since $b \neq b'$, $A \neq A'$. Under these transformations the above equation goes over into

$$(7) \qquad m^2(1 - F^{*2}) = (1 - q^{*2})Q^{*\prime}(q^*)^2.$$

This is a differential equation satisfied by the trigonometric polynomial $T_m(q^*)$. What is more, we shall show that there can be at most two polynomials of degree m which satisfy this differential equation, and that such polynomials are negatives of each other. Since $F^* = Q^*(q^*)$ satisfies the equation, this will prove that Q^* is either T_m or $-T_m$.

13. Let $H(q^*)$ be any polynomial in q^* of degree m which satisfies (7). Let a_i be the coefficient of the $(m-i)$-th power of q^* in H, i ranging from zero to m. When H and its derivative are substituted into (7) the resulting equation asserts the identity of two polynomials of degree $2m$. Let i be a non-negative integer not greater than m. The coefficient of the $2m - i$-th power of q^* in the left member is $-2m^2a_0a_i$ plus a homogeneous polynomial of degree two in the letters a_k, $k < i$, with integral coefficients. The coefficient of the same power of q^* in the right member is $-2m(m-i)a_0a_i$ plus a similar polynomial. By equating these coefficients it is possible to solve for the ratios a_i/a_0, $i = 1, \cdots, m$ as rational numbers, there being precisely one solution. This determines H up to a multiplicative constant. Reference to (7) shows that this constant is either 1 or -1. This suffices to verify our assertion.[10]

14. We now show that $F - A$ and $F - A'$ each have factors which do not divide $P'(p)$. If $F - A$ had no such factor we should have

$$F - A = (p - a_1)^{v_1+1} \cdots (p - a_k)^{v_k+1}$$

with each $v_i > 0$. The h. c. f. of $F - A$ and $P'(p)$ is a polynomial in p of degree Σv_i. The extreme case in which each v_i is unity is excluded, since then $F - A$ would be the square of a polynomial in p and $P(p)$ would be composite. Therefore $\Sigma v_i > \frac{1}{2}n$ and consequently the h. c. f. of $P'(p)$ and $F - A'$ would have a degree in p less than $n - 1 - \frac{1}{2}n = \frac{1}{2}n - 1$. It follows that $F - A'$ has at least three linear (in p) factors which do not divide $P'(p)$. Our discussion of $Q(q)$ permits us to write

$$(8) \quad \mathfrak{N}_{x|p}(F'(x)) = P'(p)^m \mathfrak{N}_{x|p}(p'(x)) = c(F - A)^{(m-1)/2}(F - A)^{(m-1)/2}\mathfrak{N}_{x|p}(q'(x))$$

---

[10] The artificial character of this argument might be avoided as follows. The coefficients of the solution of the differential equation are rational numbers. It is thus possible to use the standard theory of differential equations to solve the equation, and such a procedure would also lead to the above results.

from which we see that if $F-A$ has a linear (in p) factor which does not divide $P'(p)$ its $\frac{1}{2}(m-1)$-th power must divide $\mathfrak{N}_{x|p}(p'(x))$. Since the $\frac{1}{2}(m-1)$-th powers of three distinct linear factors cannot all be absorbed by $\mathfrak{N}_{x|p}(p'(x))$ we have our result that $F-A$ and $F-A'$ contain factors prime to $P'(p)$. The reasoning used earlier shows that $P(p)$ is equivalent to the trigonometric polynomial $T_n(p)$. In addition, we see from (8) that

$$\mathfrak{N}_{x|p}(p'(x)) = d[(p-a)(p-a')]^{(m-1)/2}$$

where $p-a$ and $p-a'$ are the factors of $F-A$ and $F-A'$ which are prime to $P'(p)$, d being some constant. One readily concludes that $p(x)$ is equivalent to the polynomial $T_m(x)$. Similarly $q(x)$ is equivalent to $T_n(x)$.

Exclusion of other possibilities.

15. We now turn to the remaining case

$$\mathfrak{N}_{p|F}(P'(p)) = c(F-A)^{v}(F-A')^{v'}, \qquad v+v'=n-1,$$

where both v and v' are different from zero, A is distinct from A' and $F-A$ has no factor $q-b$ prime to $Q'(q)$. We show that this case cannot arise; that these assumptions lead to a contradiction. This will complete our analysis of the factorization of $F'(x)$ and prove that the two decompositions treated above are the only ones possible for F.

16. We need the fact that if $p-a$ and $q-b$ both divide $F-E$, where E is any element of K, then $p-a$ and $q-b$ have at least one common factor $x-e$. To prove this, suppose $F-E=(q-b)^kF^*$, where F^* is a polynomial in q prime to $q-b$. The statement that $p-a$ divides $F-E$ and is prime to $q-b$ implies that $G=F^*/(p-a)$ is a polynomial in x. $\mathfrak{N}_{x|q}(G)$ is then a polynomial in q. Evaluating $\mathfrak{N}_{x|q}(G)$ we obtain $(F^*)^n/(F-E)$ which is obviously not a polynomial in q.

On this basis we can derive a relation between the degree in p of the h. c. f. of $F-A$ and $P'(p)$, and the degree in x of the h. c. f. of $F-A$ and $q'(x)$. Let

$$F-A=\prod_{i=1}^{g}(p-a_i)^{u_i}=\prod_{j=1}^{l}(q-b_j)^{v_j}. \tag{9}$$

We show that for each i with $u_i>1$ there is an $x-c_i$ which divides both $p-a_i$ and $q'(x)$. The multiplicity in $F-A$ of each factor $x-c_i$ which divides $p-a_i$ is a multiple of u_i. If $x-c_i$ divides $q-b_k$ its multiplicity in $F-A$ is $e_{ik}v_k$ where e_{ik} is its multiplicity in $q-b_k$. If $p-a_i$ were prime to $q'(x)$ each e_{ik} would be unity and each v_k would be divisible by u_i. Thus $F-A$ would be the u_i-th power of a polynomial in q and consequently Q

would be composite. This proves our statement. Similar remarks hold for $Q'(q)$ and $p'(x)$.

17. Since $F - A$ contains no factor prime to $Q'(q)$, the total multiplicity of the factors (in q) of $Q'(q)$ which divide $F - A$ exceeds $m/2$. Let this multiplicity be $(m/2) + j$ (where $j > 0$ need not be integral). Then

$$(10) \quad \begin{aligned} \mathfrak{N}_{x|p}(F'(x)) &= P'(p)^m \mathfrak{N}_{x|p}(p'(x)) \\ &= H(F)(F - A)^{(m/2)+j}(F - A')^k \mathfrak{N}_{x|p}(q'(x)) \end{aligned}$$

where $H(F)$ is a polynomial prime to $(F - A)(F - A')$. Then, if $p - a$ divides $F - A$ but not $P'(p)$, $(p - a)^{(m/2)+j}$ divides $\mathfrak{N}_{x}|_{p}(p'(x))$. Since $\mathfrak{N}_{x|p}(p'(x))$ is of degree $m - 1$ in p there is at most one such factor.

18. It is convenient to eliminate, first, the possibility $k \geq m/4$. From (10) we see that if h factors $p - a'$ of $F - A'$ are prime to $P'(p)$ then hk such factors must divide $\mathfrak{N}_{x|p}(p'(x))$. We now make an estimate of the degree in x of the h. c. f. of $p'(x)$ and $F - A'$, and the requirement that hk shall not exceed this degree will lead to a contradiction. Returning to (9) we see that $\Sigma(v_i - 1) = m - l$, and since the total degree of $Q'(q)$ in q is $m - 1$, we see that $m - l + k \leq m - 1$, or $l \geq k + 1$. There are, then, at least $k + 1$ distinct factors $q - b_i$ of $Q'(q)$ which divide $F - A$; there are consequently at least $k + 1$ distinct factors $x - c_i$ of $p'(x)$ which divide $F - A$. Thus the degree in x of the h. c. f. of $p'(x)$ and $F - A'$ does not exceed $m - k - 2$.

The result $hk \leq m - k - 2$ implies $k(h + 1) < m$. Recalling that $k \geq m/4$ we see that $h \leq 2$. Then there are at most two factors $p - a'$ of $F - A'$ which do not divide $P'(p)$. Under these circumstances $F - A$ must have a factor $p - a$ prime to $P'(p)$, since the contents of **14** show that if $F - A$ had no such factor $F - A'$ would have at least three.

19. This furnishes a sharper estimate for the degree in x of the h. c. f. of $p'(x)$ and $F - A'$. It cannot exceed $m - 1 - [(m/2) + j] < m/2$. Since we now know that $hk < m/2$ the requirement $k \geq m/4$ implies that h is either zero or unity. It cannot be that h is unity, for then $P(p)$ would be the trigonometric polynomial T_n, and a repetition of our earlier argument would show that $Q(q)$ is T_m, a possibility excluded by our assumptions about the factorization of $F - A$. It is also impossible that h be zero, for this would require $F - A$ to have at least three factors $p - a$ prime to $P'(p)$; whereas, in fact, it has at most one.

20. The case $k < m/4$ remains. To settle this case we observe first that since $F - A$ and $P'(p)$ have at least one factor in common it follows that $F - A$ and $q'(x)$ have at least one factor $x - c$ in common. Consequently the degree in x of the h. c. f. of $F - A'$ and $q'(x)$ does not exceed $n - 2$. It will also be useful for us to note that we may assume that n exceeds two. For, if n were two, both q and P would be equivalent to the square of a polynomial, and we should be led back to case I. This is in conflict with the assumptions made in **15**.

21. Let f be the largest integer for which $f(n-2) \leq m$ (unity is always available). Let d denote the number of distinct factors $p - a'_i$ of $F - A'$ which divide $P'(p)$. Each one appears in $\mathfrak{N}_{x|p}(F'(x))$ as follows:

$$(p - a'_i)^{mw_i}(p - a_i)^{r_i} = (p - a'_i)^{(w_i+1)k}(p - a'_i)^{t_i}.$$

This last equation and the assumption $k < m/4$ show that when $w_i \neq 0$, $t_i \neq 0$. Certainly some $t_i \leq m/fd$. From

$$mw_i \leq (w_i + 1)k + t_i \leq (w_i + 1)k + m/fd \tag{11}$$

we conclude that

$$k \geq m(w_i - (1/fd))/(w_i + 1)$$

and $k < m/4$ only if $w_i = f = d = 1$. Thus $P'(p)$ has at most one factor $p - a'$ in common with $F - A'$, and its multiplicity in $P'(p)$ must be unity. In addition, the result $f = 1$ implies $2(n-2) > m$.

22. We now show that $k = 1$. The fact that $F - A'$ has only one factor $p - a'$ in common with $P'(p)$ and that the multiplicity of $p - a'$ in $P'(p)$ is unity implies that $F - A'$ has $n - 2$ factors prime to $P'(p)$. Equation (10) reveals that $k(n-2)$ factors must divide $\mathfrak{N}_{x|p}(p'(x))$. Since the degree of this expression is $m - 1$ our result about f shows that k is unity.

We return to equation (11). If we replace w_i and k by unity, and t_i by $n - 2$ the inequality is, if anything, strengthened. We conclude that

$$m \leq 2 + (n - 2) = n.$$

This contradiction to our initial assumption about the relative size of m and n completes the proof.

University of Pennsylvania.

ON GROUPS WHOSE ORDER CONTAINS A PRIME NUMBER TO THE FIRST POWER I.*

By RICHARD BRAUER.

To the memory of I. Schur.

Introduction. Since the foundation of the theory of group characters by Frobenius in 1896, the characters of many special groups $\mathfrak{G}$ of finite order have been determined. In these examples regularities appear which, it seems, are not explained by the general theory. Their occurrence depends on certain assumptions concerning $\mathfrak{G}$; it would be extremely difficult to determine the exact nature of these assumptions and their consequences from the known examples.

In order to derive new properties of group characters we can use the theory of the modular representations of a group as a possible approach, particularly when we are interested in questions of a somewhat more arithmetical nature. The relationship between the ordinary and the modular theory can be described in the following manner: Let K be an algebraic number field. The group $\mathfrak{G}$ defines an associative algebra, the group ring Γ, and the investigation of the structure of Γ is the main subject of the ordinary theory of group characters. If we consider only elements of Γ which are linear combinations of group elements with integral coefficients we obtain an "integral domain" J in Γ which is, in general, not a maximal domain. We may study the group ring from an arithmetical point of view; in particular, we are interested in the prime ideal divisors of a fixed rational prime p. The study of this question leads to the modular theory of group characters. There is a close connection between the algebraic and arithmetical structure of Γ which implies that the theories of ordinary and modular group characters are interrelated. Thus the "modular" theory provides a new approach to the "ordinary" theory of group characters.[1]

In the case in which we are interested the prime is a discriminant divisor, i. e., p divides the order g of the group $\mathfrak{G}$. We shall assume in this paper that g contains p to the first power only, i. e.,

$$(*) \qquad g = pg', \qquad (g', p) = 1.$$

---

* Received April 4, 1941; Presented to the American Mathematical Society April 7, 1939.

[1] For this point of view, cf. R. Brauer, *Proceedings of the National Academy of Sciences*, vol. 25 (1939), p. 252.

To be sure, this assumption is of a very restrictive nature. It may, however, be mentioned that Dickson's list[2] of 78 simple groups of an order smaller than 10^9 contains only one group for which there is no prime p such that (*) holds.

The simplification imposed by the assumption (*) is threefold. In the first place, the arithmetical structure is far simpler than in the general case, and the modular theory as developed in a number of previous papers[3] gives complete results. Furthermore, the group-theoretical situation is greatly simplified compared with the general case because of the simple structure of the p-Sylow-subgroup of $\mathfrak{G}$. Finally, we have simplifications of a somewhat more analytical nature; in certain inequalities, as they occur in the theory of group characters, the terms are small so that it is easy to handle them.

Let K be the number of classes of conjugate elements in $\mathfrak{G}$. The characters of $\mathfrak{G}$ can be arranged in the form of a matrix Z of degree K. Each row corresponds to a fixed character, each column to a fixed class of conjugate elements. We arrange the columns so that the classes with p-regular elements[4] come before the other classes, and we arrange the rows so that the characters of a degree prime to p come before the characters of a degree divisible by p. Thus if

$$(**)\qquad Z = \begin{array}{c} \begin{pmatrix} Z_1 & \qquad Z_2 \\ Z_3 & \qquad Z_4 \end{pmatrix} \\ \underbrace{}_{\substack{p\text{-regular}\\ \text{classes}}} \quad \underbrace{}_{\substack{p\text{-singular}\\ \text{classes}}} \end{array} \begin{array}{l} \}\ \text{degrees prime to } p, \\ \}\ \text{degrees divisible by } p, \end{array}$$

then we have first of all that $Z_4 = 0$. Our main result states that the matrix Z_2 is determined by the structure of the normalizer $\mathfrak{M}$ of $\mathfrak{P}$, the p-Sylow-subgroup of $\mathfrak{G}$, only some $\pm$ signs remaining undetermined. In particular, if we replace $\mathfrak{G}$ by $\mathfrak{M}$, the matrix Z_2 remains essentially the same, only the signs of some of its rows have to be changed. In the case of the group $\mathfrak{M}$, the second row in (**) is missing. This implies that the number of characters of $\mathfrak{G}$ whose degree is prime to p is equal to the number of classes of conjugate elements of $\mathfrak{M}$.

We obtain these results in a somewhat more explicit form (cf. Theorems 1, 2, 3, 4, and 5). We show that the values of the characters of $\mathfrak{G}$ for p-singular elements can be written down (apart from $\pm$ signs which remain undetermined), provided that the characters of a certain subgroup $\mathfrak{W}$ are known. This $\mathfrak{W}$ is a subgroup of the normalizer $\mathfrak{M}$ of the p-Sylow-subgroup;

---

[2] L. E. Dickson, *Linear groups*, Leipzig (1901), pp. 309-310.

[3] Cf. the papers mentioned in the bibliography.

[4] Cf. the list of notations at the end of the Introduction.

the order of $\mathfrak{W}$ is prime to p, and hence the construction of its characters can certainly be considered as a more elementary problem than the corresponding problem for $\mathfrak{G}$. As a matter of fact, in most of the applications of the theory developed here the group $\mathfrak{W}$ is of a very simple structure, and its characters can be written down at once. Very often, $\mathfrak{W}$ is a cyclic group.

From a knowledge of the matrices Z_2 and Z_4 in (**), some information concerning Z_1 and Z_3 can be obtained. First, the orthogonality relations for group characters can be applied. Secondly, we may obtain the values (mod p) of the characters for those classes whose elements commute with elements of order p. In particular, this gives a set of new conditions for the degrees of the characters which in many cases is sufficient for the computation of the degrees. Finally, from our knowledge of Z_2 and Z_4 we obtain conditions for the multiplication of the characters which also can be used to gain new information concerning Z_1 and Z_3.

In many cases it is possible to derive the complete table of characters in this manner; the following examples may show the significance of the method. Heretofore it has not been known whether more than one simple group exists for the orders $g = 5616$ and $g = 6048$. Assuming that $\mathfrak{G}$ is a simple group of one of these orders, we may find by elementary methods the structure of $\mathfrak{N}$, where $p = 13$ in the first case and $p = 7$ in the second case. It turns out that the results obtained are sufficient to construct the complete table of group characters in either case. From this table we can derive the modular group characters of $\mathfrak{G}$ for the prime $p = 3$ (here the theorem as given above does not hold since 3 divides the order of the group to a higher power than the first). The values of the modular characters show that a simple group $\mathfrak{G}$ of order 5616 must have an absolutely irreducible representation of degree 3 in the Galois field $GF(3)$. Hence it is a subgroup of $LF(3,3)$, and since $\mathfrak{G}$ and $LF(3,3)$ have the same order, we find $\mathfrak{G} \cong LF(3,3)$. Similarly, it can be shown that a simple group of order 6048 is isomorphic with $HO(3,9)$ since it must have a unitary representation of degree 3 in $GF(9)$.[5] This shows[6] that the list of known simple groups is actually complete up to the order 6232. Some applications of our results of a more general nature have been given without proof elsewhere.[7]

---

[5] It is not possible to give the details of the computation in this paper.

[6] F. N. Cole, *Bulletin of the American Mathematical Society*, vol. 30 (1924), p. 489. Actually, it follows from a theorem of Frobenius *Sitzungsber. Preuss. Akad.* (1895), p. 1041) that no simple group of order 6232 can exist, since $6232 = 2 \cdot 2 \cdot 2 \cdot 19 \cdot 41$ is a product of five primes. It would be very easy to replace 6232 by a larger number using the methods of this paper.

[7] R. Brauer, *Proceedings of the National Academy of Sciences*, vol. 25 (1939), p. 290.

Notation. $\mathfrak{G}$ is a group of order $g = pg'$, where p is a prime and g' is not divisible by p. Let $\zeta_1, \zeta_2, \cdots, \zeta_K$ be the absolutely irreducible characters of $\mathfrak{G}$ and $\phi_1, \phi_2, \cdots, \phi_{k_0}$ the irreducible modular characters. Then K is the number of classes of $\mathfrak{G}$, and k_0 is the number of classes consisting of p-regular elements. Here an element G of $\mathfrak{G}$ is termed p-regular, if its order is prime to p, otherwise G is p-singular. The characters ζ_μ of $\mathfrak{G}$ are distributed into "*blocks*"[8] B_λ $(\lambda = 1, 2, \cdots)$, and we may speak of the modular characters belonging to the block B_λ. The degree of ζ_μ will be denoted by z_μ. Each character ζ_μ determines a family of p-conjugate characters,[9] we denote the number of members of this family by r_μ.

1. The p-singular elements of $\mathfrak{G}$. We first formulate some elementary facts concerning the type of groups we propose to consider in this paper.

LEMMA 1. *Let $\mathfrak{G}$ be a group of order*

$$g = pg', \qquad ((p, g') = 1), \tag{1}$$

where p is a prime, and let P be an element of order p. The normalizer $\mathfrak{N} = \mathfrak{N}(P)$ is a direct product of the Sylow subgroup $\mathfrak{P} = \{P\}$, and a group $\mathfrak{V}$ of order v which is prime to p, i. e.,

$$\mathfrak{N} = \mathfrak{V} \times \mathfrak{P}. \tag{2}$$

This follows at once from a theorem of Burnside or a more general theorem of Frobenius.[10] Using Sylow's theorem, we obtain easily

LEMMA 2. *If two elements of $\mathfrak{N}$ are conjugate in $\mathfrak{G}$, then they are conjugate in the normalizer $\mathfrak{M} = \mathfrak{N}(\mathfrak{P})$ of the Sylow-subgroup $\mathfrak{P}$.*

There is no difficulty in discussing the structure of $\mathfrak{M}$. We have

---

[8] Cf. [1], § 9.

[9] Cf. [2], § 4.

[10] W. Burnside, *Proceedings of the London Mathematical Society*, vol. 33 (1901), pp. 163, 257; G. Frobenius, *Sitzungsber. Preuss. Akad.* (1901), pp. 1216, 1324. We may also prove the statement in our case in the following fashion. Assume first that $\mathfrak{N}$ has an irreducible representation of degree f prime to p which does not represent P by the unit matrix. Then the elements of $\mathfrak{N}$ for which the determinant is a g'-th root of unity, form a normal subgroup $\mathfrak{V}$, and P does not lie in $\mathfrak{V}$, whence we easily obtain (2). If Lemma 1 is not true, then every irreducible representation $\mathfrak{F}$ of $\mathfrak{N}$, whose degree f is prime to p, actually is a representation of $\mathfrak{V}/\{P\}$. But adding the squares of the degrees of all irreducible representations, we obtain the order of the group, and it would follow that the orders of $\mathfrak{V}$ and of $\mathfrak{V}/\{P\}$ are congruent $(\mathrm{mod}\, p^2)$, which is not the case.

LEMMA 3. *The group $\mathfrak{M}/\mathfrak{N}$ is cyclic, its order q is a divisor of $p-1$, say*

$$p-1=qt. \tag{3}$$

If M in $\mathfrak{M}$ corresponds to a generating element of $\mathfrak{M}/\mathfrak{N}$, then

$$M^{-1}PM=P^{\gamma^t} \tag{4}$$

where γ is a primitive root (mod p). *The order of M is prime to p.*

The elements

$$P_0=1,\quad P_1=P,\quad P_2=P^{\gamma},\cdots,P_t=P^{\gamma^{t-1}} \tag{5}$$

form a complete system of representatives of the classes of conjugate elements in $\mathfrak{G}$ in which the order is a power of p. They can be taken as the elements P_i in [2], § 1.

Every p-singular class C of $\mathfrak{G}$ contains an element $P^{\alpha}V$ with $\alpha\not\equiv 0$ (mod p), V in $\mathfrak{V}$. We readily obtain

LEMMA 4. *Let $V_1, V_2,\cdots, V_l$ be a maximal system of elements of $\mathfrak{V}$ such that no two of them are conjugate in $\mathfrak{M}$. Every p-singular class of $\mathfrak{G}$ contains an element $P^{\alpha}V_{\lambda}$ with $\alpha\not\equiv 0$ (mod p), and λ is uniquely determined by the class. All the classes for which λ has the same value belong to one family*[11] *F_{λ}; we have l families of p-singular classes.*

Let τ_{λ} be the least positive integer for which $M^{-\tau_{\lambda}}V_{\lambda}M^{\tau_{\lambda}}$ and V_{λ} are conjugate in $\mathfrak{V}$. Then the class of V_{λ} in $\mathfrak{M}=\{\mathfrak{V},P,M\}$ splits into τ_{λ} sub-classes of elements which are conjugate in $\mathfrak{V}$. The number τ_{λ} must be a divisor of q, since M^q lies in $\mathfrak{N}$ and, therefore, $M^{-q}V_{\lambda}M^q$ and V_{λ} are conjugate in $\mathfrak{V}$.

Suppose that $P^{\alpha}V_{\lambda}$ and $P^{\beta}V_{\lambda}$ are conjugate in $\mathfrak{G}$, say $G^{-1}P^{\alpha}V_{\lambda}G=P^{\beta}V_{\lambda}$. Then $G^{-1}P^{\alpha}G=P^{\beta}$, $G^{-1}V_{\lambda}G=V_{\lambda}$. Hence G lies in one of the cosets $\mathfrak{N}M^{\rho}$. Since $M^{-\rho}V_{\lambda}M^{\rho}$ and V_{λ} are conjugate in $\mathfrak{V}$, the exponent ρ must be a multiple of τ_{λ}. Furthermore, (4) implies that $\alpha^{\gamma^{\rho t}}\equiv\beta$ (mod p). Conversely, if this congruence holds for a multiple ρ of τ_{λ}, then $P^{\alpha}V_{\lambda}$ and $P^{\beta}V_{\lambda}$ are conjugate in $\mathfrak{G}$. It follows that the elements $P^{\alpha}V_{\lambda}$, $\alpha=1,2,\cdots,p-1$, belong to $t_{\lambda}=t\tau_{\lambda}$ different classes in $\mathfrak{G}$.

LEMMA 5. *Suppose that the class of V_{λ} in $\mathfrak{G}$ breaks up into τ_{λ} sub-classes of elements which are conjugate in $\mathfrak{V}$, and set $t_{\lambda}=t\tau_{\lambda}$. The family F_{λ} consists of t_{λ} classes of conjugate elements* (*with regard to* $\mathfrak{G}$).

It follows from (3) and $\tau_{\lambda}|q$ that t_{λ} divides $p-1$. Two elements $P^{\alpha}V_{\lambda}$

---

[11] Cf. [2], § 7.

and $P^\beta V_\lambda$ are conjugate in $\mathfrak{G}$, if and only if P^α and P^β are conjugate in the normalizer $\mathfrak{N}(V_\lambda)$ of V_λ. Hence we have t_λ classes of elements of $\mathfrak{N}(V_\lambda)$ which contain elements of order p. Each of these classes contains $n(V_\lambda)/n(V_\lambda P)$ elements, where $n(G)$ denotes the order of the normalizer of an element G in $\mathfrak{G}$. Hence $\mathfrak{N}(V_\lambda)$ contains $t_\lambda n(V_\lambda)/n(V_\lambda P)$ elements of order p. Together with the unit element, they form a complete set of solutions of the equation $X^p = 1$. Using a theorem of Frobenius,[12] we now obtain

LEMMA 6. *Let $n(G)$ be the order of the normalizer of the element G in $\mathfrak{G}$. The number t_λ is characterized completely by the two conditions*

$$p - 1 \equiv 0 \pmod{t_\lambda}, \qquad 1 + t_\lambda \frac{n(V_\lambda)}{n(V_\lambda P)} \equiv 0 \pmod{p}. \tag{6}$$

2. The blocks of characters. A block B is either of type[13] 1 or of type 0, according as the degrees of the characters in B are divisible by p or are prime to p. In the first case, B is of the highest type. Therefore,[14] B consists of one character ζ_μ only. Further, ζ_μ is modular irreducible, it lies in the field of the g'-th roots of unity, and it vanishes for all p-singular elements of $\mathfrak{G}$. The corresponding generalized decomposition numbers $d^i_{\mu\nu}$ vanish for $i > 0$ and all ν.[15]

In the notation of Lemma 4, there exist l classes of conjugate elements in $\mathfrak{G}$ such that the elements of the class are p-regular and that the number of elements in the class is prime to p. This implies,[16] that exactly l of the blocks of $\mathfrak{G}$ are of the lowest type, i. e., of type 0.

THEOREM 1. *The group $\mathfrak{G}$ possesses l blocks $B_1, B_2, \cdots, B_l$ of lowest type where l is defined in Lemma 4. All the other blocks $B_{l+1}, B_{l+2}, \cdots$ are of highest type; they consist of one character ζ_μ, which is p-conjugate only to itself, which is modular irreducible, and which vanishes for all p-singular elements.*

If we set $g = p^a g'$ with $(g', p) = 1$, then $a = 1$. Hence the blocks $B_1, B_2, \cdots, B_l$ are of degree $a - 1$. Consequently, the results of [3] will hold for them. We shall apply these later.

3. The matrix of the generalized decomposition numbers. In order to discuss the decomposition numbers $d^i_{\mu\nu}$ with $i > 0$, we have to form the normalizer $\mathfrak{N}(P_i)$ of the element P_i (cf. (5)). This normalizer is identical with the normalizer $\mathfrak{N}$ of P, and (2) now shows that the irreducible modular characters $\phi_1{}^i, \phi_2{}^i, \cdots$ of $\mathfrak{N}(P_i)$ are identical with the ordinary irreducible characters $\theta_1, \theta_2, \cdots$ of $\mathfrak{V}$. The corresponding Cartan invariants of $\mathfrak{N}(P_i)$, then, are given by

---

[12] Frobenius, *Sitzungsber. Preuss. Akad.* (1895), p. 981.

[13] [1], § 17.

[14] Cf. [1], Part II.

[15] [2], Theorem 10.

[16] [1], Theorem 2.

$$c^i{}_{\rho\sigma} = p\delta_{\rho\sigma}{}^{17} \quad (\text{for } i > 0). \tag{7}$$

This implies that the indecomposable character $\Phi_\rho{}^i$ of $\mathfrak{N}(P_i)$ associated with $\phi_\rho{}^i$ is connected with $\phi_\rho{}^i$ by the equation

$$\Phi_\rho{}^i = p\phi_\rho{}^i.$$

Then the formulae [2], (1) and (2) defining the decomposition numbers $d^i{}_{\mu\nu}$ become

$$\zeta_\mu(P_iV) = \sum_\nu d^i{}_{\mu\nu}\theta_\nu(V) \qquad (V \text{ in } \mathfrak{V}, i > 0). \tag{8}$$

$$d^i{}_{\mu\nu} = (1/v) \sum_{V \text{ in } \mathfrak{V}} \zeta_\nu(P_iV)\theta_\mu(V^{-1}) \quad \text{for} \quad i > 0.^{18} \tag{9}$$

According to (7) and [2], Theorem 4, we have

$$\sum_{\mu=1}^{K} d^i{}_{\mu\nu}\bar{d}^j{}_{\mu\rho} = 0, \text{ if } i \neq j \text{ or if } \nu \neq \rho. \tag{10 a}$$

$$\sum_{\mu=1}^{K} d^i{}_{\mu\nu}\bar{d}^i{}_{\mu\nu} = p. \tag{10 b}$$

If in (10 b) the index μ ranges only over the values corresponding to a fixed block B_λ, then the sum on the left-hand side is a rational integer which is divisible by p (cf. [2], Theorem 7). It follows that for every possible combination (i, ν) with $i > 0$ there exists exactly one block B_λ of lowest kind such that $d^i{}_{\mu\nu} = 0$ for all μ not belonging to B_λ. We denote this block B_λ by $B(i, \nu)$. If two d-columns $d^i{}_{\mu\nu}$ and $d^j{}_{\mu\rho}$ are algebraically conjugate, then obviously $B(i, \nu) = B(j, \rho)$. The same block B_λ corresponds to all the members of a family of algebraically conjugate d-columns.

Every block B_λ of lowest type appears, conversely, in the form $B(i, \nu)$ for at least one combination (i, ν) with $i > 0$. If this were not so, we would have $d^i{}_{\mu\nu} = 0$ for the values μ belonging to B_λ and for all $i > 0$ and all ν. Then (8) shows that the characters ζ_μ of B_λ vanish for all p-singular elements, and this would imply[19] that ζ_μ belongs to a block of highest type and not to a block B_λ of type 0.

---

[17] Cf. [1], § 29.—Of course, $\delta_{\rho\sigma} = 0$ for $\rho \neq \sigma$ and $\delta_{\rho\rho} = 1$.

[18] We make use of the fact that v is the p-th part of the order of $\mathfrak{N}(P_i)$ and that $p\theta_\mu$ is the indecomposable character $\Phi_\mu{}^i$.

[19] Cf. [1], Theorem 10.

According to [2], Theorem 11, the number of families of algebraically conjugate d-columns is equal to the number of families of classes of conjugate elements. Each d-column with $i = 0$ forms its own family, and so does each p-regular class. Further, we have as many d-columns with $i = 0$ as we have p-regular classes. It follows that the number of families of d-columns with $i > 0$ is equal to the number of families of p-singular classes. This latter number is l (cf. Lemma 4). Hence we have as many families of algebraically conjugate d-columns with $i > 0$ as we have blocks $B_1, B_2, \cdots, B_l$ of lowest type. The mapping $d^i_{\mu\nu} \to B(i, \nu)$ necessarily defines a $(1-1)$-correspondence between families of d-columns with $i > 0$ and blocks of the lowest type. We shall say that the d-column $d^i_{\mu\nu}$ belongs to the block B_λ, if $B(i, \nu) = B_\lambda$.

Arrange now the characters $\zeta_1, \zeta_2, \cdots, \zeta_k$ of $\mathfrak{G}$ so that the characters of B_1 come first, then those of $B_2, \cdots$, then those of B_l, and then those of the blocks $B_{l+1}, B_{l+2}, \cdots$ of highest type.

On the other hand, we choose a suitable arrangement of the columns of the matrix $\boldsymbol{D}$ formed by the decomposition numbers (cf. [2], (3)). We start with the d-columns with $i = 0$, taking them in such an order that the matrix $D = D^0$ breaks up in the form [1], (28) (i. e., arranging the modular characters $\phi_1, \phi_2, \cdots$ of $\mathfrak{G}$ so that the characters of B_1 come first, then those of B_2, etc.). After the d-columns with $i = 0$ we take the columns belonging to B_1, then those of $B_2, \cdots$, finally those of B_l. From our results above it follows that $\boldsymbol{D}$ has the following form

$$\boldsymbol{D} = \left(\begin{array}{ccccc|cccc} D_1 & 0 & \cdots & 0 & 0 & Q_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 & 0 & 0 & Q_2 & \cdots & 0 \\ & & & & & & & & \\ 0 & 0 & \cdots & D_l & 0 & 0 & 0 & \cdots & Q_l \\ 0 & 0 & \cdots & 0 & I_m & 0 & 0 & \cdots & 0 \end{array}\right). \tag{11}$$

Here, m is the number of characters ζ_μ of a degree $z_\mu \equiv 0 \pmod{p}$ and I_m denotes the unit matrix of degree m. The part of $\boldsymbol{D}$ to the left of the vertical line contains the ordinary decomposition numbers $d_{\mu\nu} = d^0_{\mu\nu}$, i. e., the matrix $D = D^0$. The columns of each fixed Q_λ are algebraically conjugate. Since $\boldsymbol{D}$ is non-degenerate, it follows that these columns are linearly independent.

4. The number of characters in the blocks of lowest kind. Denote the number of columns of Q_λ by t'_λ; the corresponding columns of $\boldsymbol{D}$ form a family of d-columns $(\lambda = 1, 2, \cdots, l)$. Besides these l families of d-columns, we have k_0 further families of d-columns $d^i_{\mu\nu}$ with $i = 0$ where k_0 is the number of classes of p-regular elements in $\mathfrak{G}$. Each of these families consists of one column. Hence the number of members in the different families of d-columns are

$$(12) \qquad t'_1, t'_2, \cdots, t'_l, \underbrace{1, \cdots, 1}_{k_0}$$

respectively.

We now apply [2], Theorem 12.[20] We have $l + k_0$ families of p-conjugate characters, and if the number of members of these families are

$$(13) \qquad r_1, r_2, \cdots, r_{l+k_0}$$

respectively, then the sets of numbers (12) and (13) coincide. Further, the classes of conjugate elements are also distributed into families. According to Lemma 4, we have l families containing p-singular classes, and the λ-th family contains t_λ classes (Lemma 5). Each p-regular class forms a family by itself; there are k_0 such families. Hence the number of members of the different families is

$$(14) \qquad t_1, t_2, \cdots, t_l, \underbrace{1, \cdots, 1}_{k_0}$$

respectively. Again, [2], Theorem 12, shows that the system (14) coincides with each of the systems (12) and (13). In particular, after changing the order of the l elements $V_1, V_2, \cdots, V_l$ if necessary, we may assume that

$$(15) \qquad t_\lambda = t'_\lambda, \qquad (\lambda = 1, 2, \cdots, l).$$

Let ϵ be a primitive g-th root of unity, and choose a rational integer γ such that γ is a primitive root (mod p) and $\gamma \equiv 1 \pmod{g'}$. The substitution

$$(16) \qquad T: \ \epsilon \rightarrow \epsilon^\gamma$$

transforms Q_λ (cf. (11)) into a matrix $Q_\lambda{}^T$. There are two ways of obtaining $Q_\lambda{}^T$ from Q_λ: (1) the substitution T permutes the d-columns and this induces a permutation of the columns of Q_λ. The coefficients of Q_λ belong to the field of the p-th root of unity and the t_λ columns of Q_λ are algebraically conjugate. It follows easily that $Q_\lambda{}^T$ can be obtained from Q_λ by a cyclic permutation of the columns.[21] The substitution T permutes the characters $\zeta_1, \zeta_2, \cdots$; actually, T induces a cyclic permutation of the characters of each family. This shows that $Q_\lambda{}^T$ can be obtained from Q_λ by a certain permutation of the rows. The assumptions of [2], Lemma 2, are satisfied since the columns of Q_λ are linearly independent as observed above. It follows that the block B_λ must contain at least one character ζ_μ such that the number r_μ of members of the family of ζ_μ is divisible by t_λ. But the systems (14) and (15) both consist of the same numbers. We readily see that the block B_λ contains a family consisting of exactly t_λ characters; all the other families belonging to B_λ consist of one character $(\lambda = 1, 2, \cdots, l)$. If B_λ contains w_λ families of characters and if $r_\alpha, r_\beta, \cdots, r_\kappa$, respectively, are the numbers of members of these w_λ families, then one of the numbers, say r_α, is equal to t_λ whereas the other $w_\lambda - 1$ numbers are equal to 1, i. e., $r_\beta = \cdots = r_\kappa = 1$. On the other hand, we have [22]

---

[20] For the case $p = 2$, cf. footnote [21] of [2].

[21] According to [2], § 1, the generalized decomposition numbers lie in the field of the p-th roots of unity, since we have $a = 1$.

$$r_\alpha r_\beta \cdots r_\kappa (1/r_\alpha + 1/r_\beta + \cdots + 1/r_\kappa) = p.$$

This now gives

$$t_\lambda(1/t_\lambda + (w_\lambda - 1)) = p,$$

$$w_\lambda - 1 = (p-1)/t_\lambda. \tag{17}$$

If a character ζ_μ is p-conjugate only to itself, i. e., $r_\mu = 1$, then ζ_μ lies in the field of the g'-th roots of unity. Hence

THEOREM 2. *Let $t_1, t_2, \cdots, t_l$ have the same significance as in Lemmas 5 and 6. If these l numbers are taken in a suitable order, then the block B_λ of lowest kind consists (a) of one family of t_λ p-conjugate characters, the "exceptional family" of B_λ and (b) $(p-1)/t_\lambda$ further characters ζ_μ which belong to the field of the g'-th roots of unity, i. e., each such ζ_μ is p-conjugate only to itself $(\lambda = 1, 2, \cdots, l)$.*

On account of [**3**], § 7, Corollaries 1 and 6, and [**3**], Theorem 5, we also have

THEOREM 3. *The block B_λ contains $(p-1)/t_\lambda$ modular characters $(\lambda = 1, 2, \cdots, l)$. The structure of B_λ is characterized by a tree T_λ with $1 + (p-1)/t_\lambda$ vertices and with $(p-1)/t_\lambda$ edges. Each of these edges E_ν corresponds to a modular character ϕ_ν of B_λ; each of the vertices V_μ corresponds to a family of p-conjugate characters. A modular character ϕ_ν of B_λ appears as a modular constituent of an ordinary character ζ_μ of B_λ, if the edge E_ν contains the vertex corresponding to ζ_μ; the multiplicity of ϕ_ν in ζ_μ then is* 1.

We may describe the tree T_λ in the customary manner [23] by a matrix with $1 + (p-1)/t_\lambda$ rows and $(p-1)/t_\lambda$ columns. Each row corresponds to a vertex, and each column corresponds to an edge. The coefficient in the i-th row and the j-th column is 1, if the j-th edge contains the i-th vertex.

---

[22] Cf. [3], Corollary 5 in § 7.

[23] Cf. O. Veblen, *Analysis situs*, second ed., p. 12.

In the other case, the coefficient is 0. Let $\bar{D}_\lambda$ be this matrix. Suppose that the numbering of the vertices is such that the last vertex corresponds to the family with t_λ members. Then the matrix D_λ in (11) is obtained from $\bar{D}_\lambda$ by adding $t_\lambda - 1$ further rows all of which are equal to the last row of $\bar{D}_\lambda$. Since $C_\lambda = D'_\lambda D_\lambda$, we can find the ordinary decomposition numbers and the Cartan invariants corresponding to the block B_λ, if we know the tree T_λ, the number t_λ and the vertex corresponding to the family with t_λ members.

5. The coefficients of the matrix Q_λ. Consider a fixed block B_λ $(\lambda = 1, 2, \cdots, l)$. The corresponding matrix Q_λ in (11) has t_λ columns, (cf. (15)), which are algebraically conjugate. The substitution (16) induces a cyclic permutation of these columns; we may arrange the columns so that the first column is carried into the second, the second into the third etc., finally the last column into the first.

As we have seen, the block B_λ contains $(p-1)/t_\lambda$ characters ζ_μ which belong to the field of the g'-th roots of unity. It follows from (9) that the corresponding numbers $d^{1}_{\mu\nu}$ lie in the same field. On the other hand, the generalized decomposition numbers lie in the field of the p-th roots of unity.[24] Consequently, the numbers $d^{1}_{\mu\nu}$ with the fixed first suffix μ are rational integers, therefore $d^{1}_{\mu\nu}$ is algebraically conjugate only with itself. The formula (11) now shows that the row of Q_λ, which corresponds to ζ_μ contains t_λ equal coefficients.

Set $h = (p-1)/t_\lambda$ and arrange the characters ζ_μ of B_λ so that the h characters lying in the field of the g'-th roots of unity are taken first and are followed by the t_λ characters of the exceptional family of B_λ. If the order of these last t_λ characters is chosen suitably, the matrix Q_λ will have the form

$$Q_\lambda = \begin{pmatrix} a_1 & a_1 & \cdots & a_1 \\ a_2 & a_2 & \cdots & a_2 \\ \cdot & \cdot & \cdot & \cdot \\ a_h & a_h & \cdots & a_h \\ \alpha & \alpha' & \cdots & \alpha^{(t_\lambda-1)} \\ \alpha' & \alpha'' & \cdots & \alpha \\ \cdot & \cdot & \cdot & \cdot \\ \alpha^{(t_\lambda-1)} & \alpha & \cdots & \alpha^{(t_\lambda-2)} \end{pmatrix}. \tag{18}$$

Here $a_1, a_2, \cdots, a_h$ are rational integers, while α is an integer of the field of the p-th roots of unity which has exactly t_λ conjugates $\alpha, \alpha', \cdots, \alpha^{(t_\lambda-1)}$ with regard to the field P of rational numbers.

---

[24] Cf. [2], § 1.

The relations (10a) and (10b) can be used to determine the coefficients of Q_λ. Because of the form (11) of $\boldsymbol{D}$, we obtain from (10a) for $j=0$ the matrix equation

$$(a_1, a_2, \cdots, a_h, \alpha, \alpha', \cdots, \alpha^{(t_\lambda-1)})D_\lambda = 0. \tag{19}$$

Secondly, (10a) applied to two suitable columns of $\boldsymbol{D}$ passing through Q_λ yields

$$\sum_{\nu=1}^{h} a_\nu^2 + \operatorname{tr}(\bar{\alpha}\alpha^{(i)}) = 0 \qquad (i = 1, 2, \cdots, t_\lambda - 1), \tag{20}$$

where $\operatorname{tr}(\omega)$ denotes the trace of a number ω of the field $\mathrm{P}(\alpha)$ with regard to P. Finally, the equation (10b) gives

$$\sum_{\nu=1}^{h} a_\nu^2 + \operatorname{tr}(\alpha\bar{\alpha}) = p. \tag{21}$$

The last t_λ rows of D_λ are equal (cf. § 4). Hence (19) can be written in the form

$$(a_1, a_2, \cdots, a_h, \operatorname{tr}(\alpha))\tilde{D}_\lambda = 0, \tag{22}$$

where $\tilde{D}_\lambda$, as in § 4, is the matrix describing the tree T_λ. We set $a_{h+1} = \operatorname{tr}(\alpha)$. From (22), it follows that if the edge E_σ of T_λ is bounded by the vertices V_κ and V_ν, then

$$a_\kappa + a_\nu = 0.$$

On comparing these equations for all h edges of T_λ, we see that

$$a_i = \pm a_j, \qquad a_{h+1} = \operatorname{tr}(\alpha) \tag{23}$$

for $i, j = 1, 2, \cdots, h+1$. In (23), the $+$ sign is to be used, when the vertices V_i and V_j are connected by an even number of edges, in the other case the $-$ sign is to be taken.

We now add the equations (20) for $i = 1, 2, \cdots, t_\lambda - 1$ and (21). Thus we obtain

$$t_\lambda \sum_{\nu=1}^{h} a_\nu^2 + \operatorname{tr}(\bar{\alpha} \sum_{i=0}^{t_\lambda-1} \alpha^{(i)}) = p,$$

or

$$t_\lambda \sum_{\nu=1}^{h} a_\nu^2 + \operatorname{tr}(\bar{\alpha}) \operatorname{tr}(\alpha) = p. \tag{24}$$

But $\operatorname{tr}(\alpha) = \operatorname{tr}(\alpha)$, and by (23), (24) takes the form

$$ht_\lambda \operatorname{tr}(\alpha)^2 + \operatorname{tr}(\alpha)^2 = p.$$

This shows that the rational integer tr (α) must have the value ± 1:

$$\text{(25)} \qquad \operatorname{tr}(\alpha) = \pm 1.$$

Let ϵ_0 be a primitive p-th root of unity and set

$$\text{(26)} \qquad \epsilon_\nu = \epsilon_0{}^{\gamma^\nu} \qquad (\nu = 0, 1, 2, \cdots, p-2),$$

where γ is a primitive root (mod p). The integer α belongs to the field $\mathrm{P}(\epsilon_0)$ and can therefore be written in the form

$$\alpha = b_0\epsilon_0 + b_1\epsilon_1 + \cdots + b_{p-2}\epsilon_{p-2},$$

where the b_ν are rational integers which are uniquely determined by α. Since α has exactly t_λ conjugates with regard to P, the substitution T^{t_λ} (cf. (16)) of period $h = (p-1)/t_\lambda$ must leave α invariant. Hence, $b_\nu = b_{\nu+t_\lambda}$. If η_ρ denotes the "Gauss period"

$$\text{(27)} \qquad \eta_\rho = \epsilon_\rho + \epsilon_{\rho+t_\lambda} + \cdots + \epsilon_{\rho+(h-1)t_\lambda}, \qquad (\rho = 0, 1, \cdots, t_\lambda - 1),$$

then

$$\text{(28)} \qquad \alpha = b_0\eta_0 + b_1\eta_1 + \cdots + b_{t_\lambda-1}\eta_{t_\lambda-1}.$$

As is easily seen, we have

$$\operatorname{tr}(\eta_\rho) = -1,$$
$$\operatorname{tr}(\eta_\rho\bar{\eta}_\sigma) = -h \text{ for } \rho \neq \sigma; \ \operatorname{tr}(\eta_\rho\bar{\eta}_\rho) = p - h.$$

On substituting the value of α from (28) in (25) and (21) and using (23) and (25), we find

$$\text{(29)} \qquad \sum_{i=0}^{t_\lambda-1} b_i = \pm 1,$$

$$\text{(30)} \qquad h + (-h)\sum_{i,j=0}^{t_\lambda-1} b_ib_j + p\sum_{i=0}^{t_\lambda-1} b_i{}^2 = p.$$

Because of (29), the first two terms on the left hand side of (30) cancel; it follows that $\Sigma b_i{}^2 = 1$, and, since the b_i are rational integers, only one of them can be different from 0. The value of this particular b_ρ then is ± 1, and we have $\alpha = \pm\eta_\rho$. After permuting the rows of Q_λ, if necessary, we may assume that $\rho = 0$, i. e.,

$$\text{(31)} \qquad \alpha = \pm\eta_0.$$

With each character ζ_μ of B_λ we associate a sign $\delta_\mu = \pm 1$; if ζ_μ appears in the i-th row of Q_λ and $i \leqq h$, set $\delta_\mu = a_i$. If i has one of the values $h+1, \cdots, h+t_\lambda$, set $\delta_\mu = a_{h+1}$. The $\pm$ sign in (31) then is $-\delta_\mu$, if ζ_μ belongs to the family with t_λ members, i. e.,

$$\text{(31*)} \qquad \alpha = -\delta_\mu\eta_0.$$

The formulae (23), (31*) determine the coefficients of Q_λ; only one ± sign remains undetermined, if the tree T_λ is known. If T_λ is not known, all the signs δ_μ remain undetermined.

6. Relations between the characters of $\mathfrak{G}$ and of $\mathfrak{B}$. It remains to determine which d-columns are algebraically conjugate to a given d-column $d^i_{\mu\nu}$, ($i > 0$, ν fixed).[25] In order to obtain these conjugate columns, we have to apply the powers of the substitution T defined in (16). The effect on (9) is that P_i is replaced by its powers $\neq 1$. This shows that for one of these conjugate columns the upper index is 1. There is no restriction in assuming that in the given column we have $i = 1$, and that the first column of each Q_λ belongs to the upper index $i = 1$. We denote the given d-column by $\mathfrak{a}$; its μ-th coefficient is

$$(32) \qquad (\mathfrak{a})_\mu = d^1_{\mu\nu} = (1/v) \sum_{V \text{ in } \mathfrak{B}} \zeta_\mu(PV)\theta_\nu(V^{-1});$$

where ν has a fixed value. Let $\mathfrak{a}^{(\rho)}$ be the column obtained from $\mathfrak{a}$ by applying T^ρ. The effect of the substitution T consists in the replacing of P by P^γ in (32), and the μ-th coefficient of $\mathfrak{a}^{(\rho)}$ is, therefore, given by

$$(33) \qquad (\mathfrak{a}^{(\rho)})_\mu = (1/v) \sum_V \zeta_\mu(P^{\gamma^\rho}V)\theta_\nu(V^{-1}).$$

Let $\mathfrak{a}$ be the column of $\boldsymbol{D}$, (cf. (11)), which contains the first column of Q_λ. Since Q_λ has t_λ rows, which form a complete family of algebraically conjugate columns, we have

$$(34) \qquad \mathfrak{a}^{(t_\lambda)} = \mathfrak{a}; \quad \mathfrak{a}^{(\rho)} \neq \mathfrak{a} \text{ for } 0 < \rho < t_\lambda.$$

Divide the number ρ by t, (cf. Lemma 3). If

$$(35) \qquad \rho = \kappa t + (j-1), \qquad 1 \leqq j \leqq t,$$

then, by (4) and (5), we have

$$(36) \qquad P^{\gamma^\rho} = M^{-\kappa}P^{\gamma^{j-1}}M^\kappa = M^{-\kappa}P_jM^\kappa.$$

The character ζ_μ has the same value for elements which are conjugate in $\mathfrak{G}$. The equation (33) can, therefore, be written in the form

$$(\mathfrak{a}^{(\rho)})_\mu = (1/v) \sum_V \zeta_\mu(M^\kappa P^{\gamma^\rho} V M^{-\kappa})\theta_\nu(V^{-1}).$$

By (36), this becomes

$$\begin{aligned}(\mathfrak{a}^{(\rho)})_\mu &= (1/v) \sum_V \zeta_\mu(M^\kappa P^{\gamma^\rho} M^{-\kappa} \cdot M^\kappa V M^{-\kappa})\theta_\nu(V^{-1}) \\ &= (1/v) \sum_V \zeta_\mu(P_j M^\kappa V M^{-\kappa})\theta_\nu(V^{-1}).\end{aligned}$$

---

[25] For the following argument, cf. [2], Theorem 2.

The element $M^{-\kappa}VM^{\kappa}$ ranges over $\mathfrak{B}$, if V ranges over $\mathfrak{B}$. If we replace V by $M^{-\kappa}VM^{\kappa}$, we obtain

$$(\mathfrak{a}^{(\rho)})_{\mu} = (1/v) \sum_{V \text{ in } \mathfrak{B}} \zeta_{\mu}(P_j V)\theta_{\nu}(M^{-\kappa}V^{-1}M^{\kappa}).$$

The expression $\theta_{\nu}(M^{-\kappa}VM^{\kappa})$ represents an irreducible character of $\mathfrak{B}$ if V ranges over $\mathfrak{B}$ and κ has a fixed value. If we denote this character by $\theta_{\sigma}(V)$, then we finally find, on account of (8), that

$$(37) \qquad (\mathfrak{a}^{(\rho)})_{\mu} = d^{j}_{\mu\sigma}.$$

For $\rho = t_{\lambda}$, we have $\rho = t_{\lambda} = \tau_{\lambda}t$ (cf. Lemma 5), and hence, $\kappa = \tau_{\lambda}$, $j = 1$. On comparing (32) and (37) for $\rho = t_{\lambda}$ and using (34), we obtain $d^{1}_{\mu\nu} = d^{1}_{\mu\sigma}$ for every μ belonging to the block B_{λ}. This implies $\nu = \sigma$, and because of the definition of θ_{σ} and of $\kappa = \tau_{\lambda}$, this gives

$$(38) \qquad \theta_{\nu}(M^{-\tau_{\lambda}}VM^{\tau_{\lambda}}) = \theta_{\nu}(V).$$

On the other hand, if $\theta_{\nu}(M^{-\kappa}VM^{\kappa}) = \theta_{\nu}(V)$ for all V in $\mathfrak{B}$ and a fixed κ, then $\sigma = \nu$. Choose now $\rho = \kappa t$ with this value of κ, (cf. (35)). From (37) and (32) it follows that $(\mathfrak{a}^{(\rho)})_{\mu} = d^{1}_{\mu\nu}$, i. e., $\mathfrak{a}^{(\rho)} = \mathfrak{a}$. For $\rho < t_{\lambda}$, i. e., for $\kappa < t_{\lambda}/t = \tau_{\lambda}$, this contradicts (34). Hence

$$(39) \qquad \theta_{\nu}(M^{-\kappa}VM^{\kappa}) \neq \theta_{\nu}(V) \text{ for } \kappa = 1, 2, \cdots, \tau_{\lambda} - 1.$$

It follows from Lemmas 1 and 3 that the characteristic subgroup $\mathfrak{B}$ of $\mathfrak{N}$ is a normal subgroup of $\mathfrak{M}$. For any irreducible character $\theta(V)$ of $\mathfrak{B}$ and any fixed element G of $\mathfrak{M}$, the expression $\theta(G^{-1}VG)$ also represents an irreducible character of $\mathfrak{B}$, if V ranges over $\mathfrak{B}$. Two such characters $\theta(V)$ and $\theta(G^{-1}VG)$ are termed *associated characters* of $\mathfrak{B}$ with regard to $\mathfrak{M}$. All the irreducible characters of $\mathfrak{B}$ appear distributed into classes of associated characters. In order to obtain all the characters associated with $\theta(V)$, it is sufficient to take G as a power of the element M (cf. Lemma 3). The relations (38) and (39) show that there exist exactly τ_{λ} distinct characters of $\mathfrak{B}$ which are associated with θ_{ν}. We may assume that the notation has been so chosen that

$$(40) \qquad \theta_{\nu}(M^{-\kappa}VM^{\kappa}) = \theta_{\nu+\kappa}(V), \qquad (\kappa = 0, 1, 2, \cdots, \tau_{\lambda} - 1).$$

Then (37) takes the form

$$(41) \qquad (\mathfrak{a}^{(\rho)})_{\mu} = d^{j}_{\mu,\nu+\kappa} \text{ if } \rho = \kappa t + (j - 1), \quad 1 \leqq j \leqq t.$$

In other words, the row of Q_{λ} which corresponds to the characters ζ_{μ} of Q_{μ} is given by

$$(42)\qquad d^1_{\mu\nu}, d^2_{\mu\nu}, \cdots, d^t_{\mu\nu}, d^1_{\mu,\nu+1}, d^2_{\mu,\nu+1}, \cdots, d^t_{\mu,\nu+1}, \cdots, \\ d^1_{\mu,\nu+\tau_\lambda-1}, d^2_{\mu,\nu+\tau_\lambda-1}, \cdots, d^t_{\mu,\nu+\tau_\lambda-1}.$$

For every ζ_μ of B_λ, all $d^i_{\mu\nu}$ with $i > 0$ not appearing in (42) vanish as follows from the form (11) of $\boldsymbol{D}$. Hence (8) reads

$$(43)\qquad \begin{aligned} \zeta_\mu(P_i V) &= \sum_{\kappa=0}^{\tau_\lambda-1} d^i_{\mu,\nu+\kappa}\theta_{\nu+\kappa}(V), \\ \zeta_\mu(P_i V) &= \sum_{\kappa=0}^{\tau_\lambda-1} d^i_{\mu,\nu+\kappa}\theta_\nu(M^{-\kappa}VM^\kappa), \qquad (i > 0,\ V \text{ in } \mathfrak{B}). \end{aligned}$$

The row (42) of Q_λ has been determined in § 5. If ζ_μ lies in the fields of the g'-th roots of unity, then all the coefficients (42) have the same value $\delta_\mu = \pm 1$ so that (42) has the form

$$(44)\qquad \delta_\mu, \delta_\mu, \cdots, \delta_\mu.$$

If ζ_μ is a suitable one of the t_λ p-conjugate characters of B_λ, then (42) is identical with

$$(45)\qquad -\delta_\mu\eta_0, -\delta_\mu\eta_1, \cdots, -\delta_\mu\eta_{t_\lambda-1}. \qquad \text{(cf. (31*))}.$$

The index ν is determined by the fact that the first column of Q_λ in (11) is the d-column $d^1_{\mu\nu}$, and Q_λ also contains the d-columns $d^1_{\mu,\nu+1}, \cdots, d^1_{\mu,\nu+\tau_\lambda-1}$ belonging to characters θ_ρ of $\mathfrak{B}$ which are associated with θ_ν. In this manner, B_λ determines a class C_λ of associated characters of $\mathfrak{B}$. To another block $\mathrm{B}_{\lambda'}$ with $1 \leqq \lambda' \leqq l$, there must correspond a class $C_{\lambda'} \neq C_\lambda$. Every class of associated characters of $\mathfrak{B}$ necessarily appears in the form C_λ. We now change the notation and denote the character θ_ν of $\mathfrak{B}$ belonging to the first column of Q_λ by θ_λ, $\lambda = 1, 2, \cdots, l$. No two of the characters $\theta_1, \theta_2, \cdots, \theta_l$ of $\mathfrak{B}$ then are associated, but every irreducible character of $\mathfrak{B}$ is associated with one of these l characters.

Collecting the results: we substitute in (43) the values of $d^i_{\mu\nu}$ obtained by comparing (42) with (44), (45) respectively and express η_ρ by means of a primitive p-th root of unity using (26) and (27). We now obtain

THEOREM 4. *To each of the blocks B_λ of the lowest kind there corresponds an irreducible character θ_λ of $\mathfrak{B}$. No two characters $\theta_1, \theta_2, \cdots, \theta_l$ of $\mathfrak{B}$ are associated with regard to $\mathfrak{M}$, but every irreducible character of $\mathfrak{B}$ is associated with one of the characters θ_λ. There are exactly $\tau_\lambda = t_\lambda/t$ characters associated with θ_λ; they are*

$$\theta_\lambda(M^{-\kappa}VM^\kappa) \qquad (\kappa = 0, 1, 2, \cdots, \tau_\lambda - 1)$$

where V ranges over $\mathfrak{B}$; we have

$$\theta_\lambda(M^{-\tau_\lambda} V M^{\tau_\lambda}) = \theta_\lambda(V), \qquad V \text{ in } \mathfrak{B}. \tag{46}$$

If ζ_μ is a character belonging to B_λ *which is p-conjugate only to itself, then*

$$\zeta_\mu(P^\rho V) = \delta_\mu \sum_{\kappa=0}^{\tau_\lambda - 1} \theta_\lambda(M^{-\kappa} V M^\kappa), \qquad (\tau_\lambda = t_\lambda / t) \tag{47a}$$

for $\rho \not\equiv 0$ (mod p) *and V in $\mathfrak{B}$. Here $\delta_\mu = +1$ or $\delta_\mu = -1$. If ζ_μ is a character of the exceptional family of* B_λ, *we have*

$$\zeta_\mu(P^\rho V) = -\delta_\mu \sum_{\kappa=0}^{q-1} \epsilon^{\rho\gamma^{\kappa t}} \theta_\lambda(M^{-\kappa} V M^\kappa), \qquad (q = (p-1)/t), \tag{47b}$$

where $\rho \not\equiv 0$ (mod p) *and V is in $\mathfrak{B}$. Here ϵ is a suitable primitive p-th root of unity, γ is a primitive root* (mod p) *and $\delta_\mu = +1$ or $\delta_\mu = -1$.*

If the characters θ of $\mathfrak{B}$ are known, the formulae (47a) and (47b) together with Theorem 1 show that the values of the characters ζ of $\mathfrak{G}$ for all p-singular elements can be obtained, only the signs $\delta_\mu = \pm 1$ remain undetermined. If the tree T_λ corresponding to B_λ is known, then we obtain from (23) and the remark at the end of § 5

THEOREM 5. *If ζ_μ and ζ_ν are two characters of a block* B_λ *of the lowest kind, then $\delta_\mu = \delta_\nu$ if and only if the corresponding vertices of the tree* T_λ *can be joined by an even number of edges.*

Concerning the values of the characters for p-regular elements we state

THEOREM 6. *Let* B_λ *be a block of lowest type. For any p-regular elements G of $\mathfrak{G}$, we have*

$$\sum_\mu \delta_\mu \zeta_\mu(G) = 0, \tag{48}$$

where ζ_μ ranges over a complete system of characters representing the different families of B_λ *(i. e., ζ_μ ranges over the $(p-1)/t_\lambda$ characters of* B_λ *which lie in the field of the g'-th roots of unity and one of the t_λ p-conjugate characters). In particular,*

$$\sum_\mu \delta_\mu z_\mu = 0, \tag{49}$$

where z_μ is the degree of ζ_μ and μ ranges over the same values as in (48).

Proof. From the definition of the δ_μ at the end of § 5 and from (22) it follows that

$$\sum_\mu \delta_\mu d_{\mu\nu} = 0 \tag{50}$$

for $\nu = 1, 2, \cdots, k_0$ where the $d_{\mu\nu} = d^0{}_{\mu\nu}$ are the ordinary decomposition numbers of $\mathfrak{G}$ and μ ranges over the same values as in (48). On multiplying (50) by the ν-th irreducible modular character $\phi_\nu(G)$ of $\mathfrak{G}$ and adding over ν, we obtain (48). The equation (49) is derived from (48) by setting $G = 1$.

7. Corollaries. Apart from the signs δ_μ, the right hand sides in (47a) and (47b) are completely determined, if (a) the characters of $\mathfrak{V}$ are known and (b) it is known in what manner these characters are associated in the group $\{\mathfrak{V}, M\}$. This group $\{\mathfrak{V}, M\}$ has an order $v(p-1)/t$ which is prime to p, and we may consider the computation of its characters as a simpler problem than the determination of the characters of $\mathfrak{G}$. In the applications of the theory developed here, the characters of $\{\mathfrak{V}, M\}$ can usually be found without difficulty and then the answers to (a) and (b) are actually known.

In particular, the right hand sides in (47a) and (47b) are determined completely by the structure of $\mathfrak{M}$, apart from the signs δ_μ. Except for these signs, the right hand side of these equations are the same for any two groups $\mathfrak{G}$ and $\mathfrak{G}_1$, for which $\mathfrak{M} = \mathfrak{N}(\{P\})$ has the same structure. In particular, we may take for $\mathfrak{G}_1$ the group $\mathfrak{M}$ itself. Since the characters of $\mathfrak{G}$ of highest kind vanish for all p-singular elements, and since $\mathfrak{M} = \mathfrak{N}(\{P\})$ has no characters of highest kind,[26] we obtain

THEOREM 7. *Arrange the matrix Z of the group characters of $\mathfrak{G}$ in such a manner that the characters of a degree prime to p occupy the upper rows and the other characters the lower rows, further arrange the columns of Z so that the classes of the p-regular elements are to the left of the p-singular elements. If Z thus is broken up in the form*

$$Z = \begin{pmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{pmatrix} \begin{matrix} \}\text{ degrees prime to } p, \\ \}\text{ degrees divisible by } p, \end{matrix}$$
$$\underbrace{\qquad}_{p\text{-regular classes}} \quad \underbrace{\qquad}_{p\text{-singular classes}}$$

then $Z_4 = 0$. If the matrix of group characters Z^* of the group $\mathfrak{M} = \mathfrak{N}(\{P\})$ is arranged in a corresponding manner, then the lower part is missing, i. e.,

$$Z^* = (Z^*{}_1, Z^*{}_2),$$

and Z_2 can be carried into $Z^*{}_2$, if certain of its rows are multiplied by -1.

This implies

THEOREM 8. *The number of irreducible characters of $\mathfrak{G}$ of a degree*

---

[26] Cf. [1], § 29.

$z \not\equiv 0 \pmod p$ *is equal to the number of classes of conjugate elements in the group* $\mathfrak{M} - \mathfrak{N}(\{P\})$.

The characteristic roots of the matrix $\mathfrak{F}(PV)$ representing PV in a given representation $\mathfrak{F}$ of $\mathfrak{G}$ can be obtained by multiplying the roots of $\mathfrak{F}(P)$ with those of $\mathfrak{F}(V)$, taken in a suitable arrangement. Hence, for the traces of the matrices, we have

$$\operatorname{tr}(\mathfrak{F}(PV)) \equiv \operatorname{tr}(\mathfrak{F}(V)) \qquad (\operatorname{mod}(1-\epsilon)),$$

if ϵ is a primitive p-th root of unity. In particular,

$$\zeta_\mu(PV) \equiv \zeta_\mu(V) \qquad (\operatorname{mod}(1-\epsilon)).$$

Combining this with (47a), we find

$$\zeta_\mu(V) \equiv \delta_\mu \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda(M^{-\kappa}VM^\kappa) \qquad (\operatorname{mod}(1-\epsilon)),$$

if ζ_μ is p-conjugate only to itself. Since both sides here lie in the field of the g'-th root of unity, this congruence must hold $(\operatorname{mod} p)$. Similarly, we have in the case of (47b)

$$\zeta_\mu(V) \equiv -\delta_\mu \sum_{\kappa=0}^{p-1} \theta_\lambda(M^{-\kappa}VM^\kappa) \qquad (\operatorname{mod} p).$$

On the right hand side, every term $\theta_\lambda(M^{-\kappa}VM^\kappa)$ $(\kappa = 0, 1, \cdots, \tau_\lambda - 1)$ appears $(p-1)/t_\lambda$ times because of (46). Hence

THEOREM 9. *For an element V of $\mathfrak{B}$, we have* $(\tau_\lambda = t_\lambda/t)$

$$\zeta_\mu(V) \equiv \delta_\mu \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda(M^{-\kappa}VM^\kappa) \qquad (\operatorname{mod} p); \tag{51a}$$

if ζ_μ is a character of B_λ $(\lambda = 1, 2, \cdots, l)$ which is p-conjugate only to itself. If ζ_μ belongs to the exceptional family of B_λ then

$$\zeta_\mu(V) \equiv (\delta_\mu/t_\lambda) \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda(M^{-\kappa}VM^\kappa) \qquad (\operatorname{mod} p). \tag{51b}$$

Finally, if ζ_μ is a character of a degree z_μ divisible by p, we have

$$\zeta_\mu(V) \equiv 0 \qquad (\operatorname{mod} p). \tag{51c}$$

As a corollary, we obtain

THEOREM 10. *Let f_λ be the degree of the character θ_λ of $\mathfrak{B}$. The degree z_μ of a character ζ_μ of the block B_λ $(\lambda = 1, 2, \cdots, l)$ satisfies the congruence* $(\operatorname{mod} p)$

(52a) $z_\mu \equiv \delta_\mu t_\lambda f_\lambda / t$ *or* (52b) $z_\mu \equiv \delta_\mu f_\lambda / t$ $\pmod{p}$,

according as we have the case of (51a) *or* (51b).

These congruences, together with (49) and the fact that z_μ divides g are in many cases sufficient to determine the degrees. If the degrees z_μ and the degrees f_λ are known, then the sign δ_μ can be obtained from (52a), (52b), at least for an odd p.

8. The block of the 1-character. We consider now the 1-character of $\mathfrak{G}$. We may assume that the notation is so chosen that it is the character $\zeta_1(G)$, i. e., $\zeta_1(G) = 1$ for every G in $\mathfrak{G}$, and that it belongs to the block B_1. The formula (47a) then gives

$$1 = \delta_1 \sum_{\kappa=0}^{\tau_1-1} \theta_1(M^{-\kappa} V M^{\kappa})$$

for every V in $\mathfrak{V}$, and since this is a linear relation between the characters of $\mathfrak{V}$, we must have $\tau_1 = 1$, and $\theta_1(V) = 1$. Hence

THEOREM 11. *If the block* B_1 *contains the* 1-*character* ζ_1 *of* $\mathfrak{G}$, *then* θ_1 *is the* 1-*character of* $\mathfrak{V}$. *Further* $t_1 = t$. *The degrees* z_μ *of the characters of* B_1 *satisfy the congruences*

(53) $z_\mu \equiv \delta_\mu = \pm 1 \pmod{p}$ *or* $z_\mu \equiv \delta_\mu / t = \pm 1/t \pmod{p}$

according as we have the case of (51a) *or* (51b). *If* $\zeta_1, \zeta_2, \cdots, \zeta_{q+1}$, $(q = (p-1)/t)$, *represent the different families of* B_1, *then*

(54) $$1 + \delta_2 z_2 + \cdots + \delta_{q+1} z_{q+1} = 0.$$

UNIVERSITY OF TORONTO.

BIBLIOGRAPHY.

[1] R. Brauer–C. Nesbitt, "On the modular characters of groups," *Annals of Mathematics*, vol. 42 (1941), pp. 556-590.

[2] R. Brauer, "On the connection between the ordinary and the modular characters of groups of finite order," *Annals of Mathematics*, vol. 42 (1941), pp. 926-935.

[3] ———, "Investigations on group characters," *Annals of Mathematics*, vol. 42 (1941), pp. 936-958.

ON GROUPS WHOSE ORDER CONTAINS A PRIME NUMBER TO THE FIRST POWER II.*

By Richard Brauer.

Introduction. This paper is a continuation of a previous paper with the same title.[1] Its aim is the proof of the following theorem: Let $\mathfrak{G}$ be a finite group of linear transformations in n variables. Assume that the order g of $\mathfrak{G}$ contains a prime factor p to the first power only, and that $\mathfrak{G}$ has no normal subgroup of order p. Then we have $p \leqq 2n+1$. The equality sign can hold only when $\mathfrak{G}$, considered as a collineation group, is isomorphic with $LF(2,p)$. If the group $\mathfrak{G}$ of order $g = pg'$, with $(g',p) = 1$, has a normal subgroup $\mathfrak{P}$ of order p, then $\mathfrak{G}$ contains a normal subgroup $\mathfrak{N}$, such that $\mathfrak{N}$ is the direct product of $\mathfrak{P}$ and a normal subgroup $\mathfrak{V}$ of $\mathfrak{G}$, while the factor group $\mathfrak{G}/\mathfrak{N}$ is cyclic.[2] For a primitive group $\mathfrak{G}$ of this type, we can take $\mathfrak{G} = \mathfrak{N}$, i. e. $\mathfrak{G} = \mathfrak{P} \times \mathfrak{V}$; if $\mathfrak{G}$ is primitive and unimodular,[3] then the degree n must be divisible by p.

It has been proved by H. F. Blichfeldt[4] that the order g of a primitive unimodular linear group $\mathfrak{G}$ in n variables is not divisible by a prime number p which is greater than $(2n+1)(n-1)$. Our theorem improves this result for primes p which divide g to the first power only. Since $LF(2,p)$ has an irreducible representation by collineations in $(p-1)/2$ homogeneous variables,[5] the inequality $p \leqq 2n+1$ cannot be improved further.

The proof of the theorem is obtained by combining a rough estimate for the degrees of the characters (§ 3) with a formula for the product of certain characters (§ 4); both these formulae are derived from the results of [4]. The case $p = 2n+1$ requires a somewhat complicated discussion. For the notation employed cf. [4].

---

* Received June 26, 1941.

[1] Cf. the paper [4] of the bibliography.

[2] Cf. [4], § 1.

[3] This means that all the matrices of $\mathfrak{G}$ have determinant 1.

[4] H. F. Blichfeldt, *Finite Collineation Groups*, University of Chicago Press, Chicago, 1917, p. 89, Theorem 5.

[5] For $p \equiv -1 \pmod 4$, this representation has been discovered by F. Klein, *Mathematische Annalen*, vol. 15 (1879), p. 275; for $p \equiv 1 \pmod 4$ by I. Schur, *Journ. f. d. reine u. angew. Math.*, vol. 132 (1907), p. 135.

1. Remarks on characters. Let $\mathfrak{G}$ be a group of finite order. As is well known, the linear combinations

$$\xi = a_1\zeta_1 + a_2\zeta_2 + \cdots + a_l\zeta_l \qquad (a_\lambda \text{ rational integers})$$

of the (ordinary) irreducible characters $\zeta_1, \zeta_2, \cdots, \zeta_l$ of $\mathfrak{G}$ form a ring $\mathfrak{T}$. If $\mathfrak{T}_0$ denotes the subset of elements ξ for which all $a_\lambda \geqq 0$, then $\mathfrak{T}_0$ is closed under addition and multiplication; the elements of $\mathfrak{T}_0$ are the (reducible and irreducible) characters of $\mathfrak{G}$. Every element ξ of $\mathfrak{T}_0$ can be obtained by addition from the irreducible characters $\zeta_1, \zeta_2, \cdots, \zeta_l$, and this additive representation of ξ is unique. Every element of $\mathfrak{T}$ is a difference of two elements of $\mathfrak{T}_0$. It may be remarked that we have a certain resemblance to the familiar type of arithmetic; however, the roles of addition and multiplication are interchanged. We shall say that an element ξ of $\mathfrak{T}$ contains an element ξ_1 of $\mathfrak{T}$, if $\xi - \xi_1$ belong to $\mathfrak{T}_0$. We then write: $\xi \supseteqq \xi_1$. On using the orthogonality relations for characters, we easily obtain the following:

Lemma 1. *Let ζ_κ, ζ_λ, ζ_μ be irreducible characters of $\mathfrak{G}$. If $\zeta_\kappa\zeta_\lambda \supseteqq h\zeta_\mu$, where $h > 0$ is a rational integer, then $\bar{\zeta}_\kappa\zeta_\mu \supseteqq h\zeta_\lambda$, $\bar{\zeta}_\kappa$ denoting the conjugate complex character of ζ_κ.*

Indeed, both relations are equivalent to the fact that $\zeta_\kappa\zeta_\lambda\bar{\zeta}_\mu$ contains h times the 1-character $\zeta_1 = 1$.

2. Summary of the results of [4]. Let $\mathfrak{G}$ be a group of order

$$g = pg', \qquad (p, g') = 1,$$

where p is a prime number. Let P be an element of order p. Its normalizer $\mathfrak{N} = \mathfrak{N}(P)$ is of the form

$$\mathfrak{N} = \mathfrak{B} \times \{P\};$$

the normalizer $\mathfrak{M}$ of the p-Sylow-subgroup $\mathfrak{P} = \{P\}$ contains both $\mathfrak{N}$ and $\mathfrak{B}$ as normal subgroups. The factor group $\mathfrak{M}/\mathfrak{N}$ is cyclic. If M in $\mathfrak{M}$ corresponds to a generating element, then

$$M^{-1}PM = P^{\gamma^t}; \tag{1}$$

where γ is a primitive root (mod p), and $(p-1)/t = q$ is the order of $\mathfrak{M}/\mathfrak{N}$.

Let $\zeta_1 = 1, \zeta_2, \zeta_3, \cdots$ be the (ordinary) irreducible characters of $\mathfrak{G}$, and denote by $z_\mu = Dg(\zeta_\mu)$ the degree of ζ_μ and by r_μ the number of p-conjugate characters. These characters appear distributed into blocks $B_1, B_2, \cdots$ of characters. A block B_λ is either of the type 1 (highest type) or of type 0. In the first case, B_λ consists of exactly one character ζ_μ; we have $z_\mu \equiv 0$ (mod p) and $r_\mu = 1$. In the second case, all the characters of B_λ have degrees which are relatively prime to p. Let $B_1, B_2, \cdots, B_l$ be the blocks of type 0. To each of these l blocks, there corresponds a certain multiple $t_\lambda > 0$ of t. The block B_λ then consists of $q_\lambda = (p-1)/t_\lambda$ characters ζ_μ for which $r_\mu = 1$, and one "exceptional" family of t_λ p-conjugate characters.[6] Set $t_\lambda/t = \tau_\lambda$. To each block B_λ of type 0 there corresponds a class of irreducible characters of $\mathfrak{B}$, say

$$\theta_\lambda(V), \theta_\lambda^{(1)}(V) = \theta_\lambda(M^{-1}VM), \cdots, \theta_\lambda^{(\tau_\lambda-1)}(V) = \theta_\lambda(M^{-(\tau_\lambda-1)}VM^{\tau_\lambda-1}),$$

which are associated in $\mathfrak{M}$; we have

$$\theta_\lambda(M^{-\tau_\lambda}VM^{\tau_\lambda}) = \theta_\lambda(V).$$

Each irreducible character of $\mathfrak{B}$ appears exactly once in the form

$$\theta_\lambda^{(\kappa)} \qquad (\lambda = 1, 2, \cdots, l;\ \kappa = 0, 1, 2, \cdots, \tau_\lambda - 1).$$

For each ζ_μ belonging to a block B_λ of type 0 a sign $\delta_\mu = \pm 1$ is defined such that the value of ζ_μ for p-singular elements $G = P^iV$ ($i \not\equiv 0 \pmod p$, V in $\mathfrak{B}$) can be obtained in the following forms:

CASE I: $r_\mu = 1$, $\delta_\mu = 1$, $z_\mu \not\equiv 0 \pmod p$

$$\zeta_\mu(P^iV) = \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)}(V); \tag{2, I}$$

CASE II. $r_\mu = 1$, $\delta_\mu = -1$; $z_\mu \not\equiv 0 \pmod p$

$$\zeta_\mu(P^iV) = -\sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)}(V); \tag{2, II}$$

CASE III. $r_\mu = t_\lambda > 1$, $\delta_\mu = 1$; $z_\mu \not\equiv 0 \pmod p$

$$\zeta_\mu(P^iV) = -\sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)}(V) \sum_\sigma \epsilon^{i\gamma^{t\sigma}}, \tag{2, III}$$

where ϵ is a suitable primitive p-th root of unity, and σ ranges over all the values with $\sigma \equiv \kappa \pmod{\tau_\lambda}$, $0 \leqq \sigma < q$;

CASE IV. $r_\mu = t_\lambda > 1$, $\delta_\mu = -1$; $z_\mu \not\equiv 0 \pmod p$

$$\zeta_\mu(P^iV) = \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)}(V) \sum_\sigma \epsilon^{i\gamma^{t\sigma}}, \tag{2, IV}$$

where ϵ, σ have the same significance as in Case III.

---

[6] If $t_\lambda = 1$, then B_λ consists of $1 + (p-1)t_\lambda = p$ characters ζ_μ, and for all of them we have $r_\mu = 1$. An arbitrary one of these characters can be selected as the only member of the exceptional family.

Finally, for characters ζ_μ belonging to a block of type 1, we have

CASE V. $z_\mu \equiv 0 \pmod{p}$, $r_\mu = 1$

$$\zeta_\mu(P^i V) = 0. \tag{2, V}$$

For each irreducible character, we have one of these five cases.

We choose B_1 as the block which contains the 1-character $\zeta_1 = 1$. Then θ_1 is the 1-character of $\mathfrak{V}$, $\theta_1 = 1$, and $t_1 = t$.

3. The characters of $\mathfrak{N}$ induced by the characters of $\mathfrak{G}$. Every character $\zeta_\mu(G)$ defines a character of $\mathfrak{N}$ obtained by restricting G to elements of $\mathfrak{N}$. The formulae (2) in § 1 are not quite sufficient to express this "induced" character by means of the irreducible characters of $\mathfrak{N}$, because we had to assume $i \not\equiv 0 \pmod{p}$ in (2). However, the formulae yield at least some information concerning this question.

Let ϵ be a primitive p-th root of unity and denote by (ϵ) the character of $\mathfrak{P} = \{P\}$ belonging to the representation $P^j \rightarrow \epsilon^j$. Every irreducible character of $\mathfrak{N} = \mathfrak{V} \times \{P\}$ is of the form

$$\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu, \tag{3}$$

$(\lambda = 1, 2, \cdots, l;\ \kappa = 0, 1, \cdots, \tau_\lambda - 1;\ \nu = 0, 1, \cdots, p-1)$. If we restrict ourselves to elements VP^j of $\mathfrak{N}$, then every ζ_μ is a linear combination of the characters (3), the coefficients being rational integers $a(\mu, \lambda, \kappa, \nu)$, i. e.

$$\zeta_\mu(VP^j) = \sum_{\lambda,\kappa,\nu} a(\mu, \lambda, \kappa, \nu)\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu. \tag{4}$$

According to the choice of ζ_μ, we have one of the cases I, II, III, IV, V, § 2. In the corresponding equation (2), set $i = 1$ and replace ϵ by (ϵ). On subtracting the expression thus obtained from (4) we find a linear combination of the characters $\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu$ which vanishes for every element VP^j with $j \not\equiv 0 \pmod{p}$. As is easily seen, the p characters $\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu$ with fixed κ and λ and $\nu = 0, 1, 2, \cdots, p-1$ must appear with the same coefficient in such a linear combination. Since the coefficients $a(\mu, \lambda, \kappa, \nu)$ in (4) are not negative, we obtain easily the following theorem:

THEOREM 1. *When we restrict ourselves to elements G of $\mathfrak{N}$, the irreducible characters $\zeta_\mu(G)$ of G belonging to the block B_λ can be expressed by the irreducible characters of $\mathfrak{N}$ in the following form.*

CASE I. $$\zeta_\mu = \sum_{\kappa=0}^{\tau_\lambda - 1} \theta_\lambda^{(\kappa)} + S;$$

CASE II. $$\zeta_\mu = \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)} \sum_{\rho=1}^{p-1} (\epsilon)^\rho + S;$$

CASE III.[7] $$\zeta_\mu = \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)} \sum_{\rho}{}' (\epsilon)^\rho + S,$$

where ρ *ranges over the values* $0 \leqq \rho < p$ *for which the* q_λ*-th power of* ρ *is not congruent* $(\operatorname{mod} p)$ *to the* q_λ*-th power of* $\gamma^{t\kappa}$, $(q_\lambda = (p-1)/t_\lambda)$;

CASE IV.[7] $$\zeta_\mu = \sum_{\kappa=0}^{\tau_\lambda-1} \theta_\lambda^{(\kappa)} \sum_{\rho}{}'' (\epsilon)^\rho + S,$$

where ρ *ranges over the values* $0 \leqq \rho < p$ *for which the* q_λ*-th power of* ρ *is congruent* $(\operatorname{mod} p)$ *to the* q_λ*-th power* $\gamma^{t\kappa}$;

CASE V. $$\zeta_\mu = S.$$

Here, (ϵ) *denotes a character of the cyclic group* $\{P\}$ *which is not the* 1*-character. The expression* S *is a linear combination of the characters* $\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu$ *of* $\mathfrak{N}$ *with non-negative integral rational coefficients such that* $\theta_\lambda^{(\kappa)}(V)(\epsilon)^\nu$ *and* $\theta_\lambda^{(\kappa)}(V) \cdot 1 = \theta_\lambda^{(\kappa)}$ *always have the same coefficient.*

As a corollary, we have

COROLLARY 1. *If* θ_λ *has the degree* f_λ*, then the degree* z_μ *of* ζ_μ *has the following forms*:

CASE I. $z_\mu = f_\lambda t_\lambda / t + ps$;

CASE II. $z_\mu = f_\lambda t_\lambda (p-1)/t + ps$;

CASE III. $z_\mu = f_\lambda t_\lambda t^{-1}(p-(p-1)/t_\lambda) + ps$;

CASE IV. $z_\mu = (p-1) f_\lambda / t + ps$;

CASE V. $z_\mu = ps$.

where $s \geqq 0$ *is a rational integer, and* $s \geqq 1$ *in* Case V.

If the degree z_μ is smaller than p, the expression S in Theorem 1 must vanish. In Case I, this implies that $\zeta_\mu(P) = z_\mu$, and therefore P is represented by the unit matrix. In Case II, we must have $z_\mu = (p-1)$, $f_\lambda = 1$, $t_\lambda = t$. Then every element V of $\mathfrak{B}$ is represented by a scalar multiple of the unit matrix: $V \rightarrow \theta_\lambda(V) \cdot I$. In Case III, we must have $f_\lambda = 1$, $t_\lambda = t$, $z_\mu = p - (p-1)/t \geqq (p+1)/2$, since $t_\lambda \geqq 2$. Again, every element of $\mathfrak{B}$ is represented by a scalar multiple $\theta_\lambda(V)$ of the unit matrix. In Case IV, we must have $z_\mu = (p-1) f_\lambda / t$; Case V is not possible. Hence

---

[7] If the character (ϵ) is chosen in a fixed manner, then this formula holds for a character chosen suitably from the exceptional family of B_λ.

COROLLARY 2. *If the degree z_μ of the irreducible representation $\mathfrak{Z}_\mu$ of $\mathfrak{G}$ is smaller than p, then we have one of the following cases:*

CASE I. $z_\mu = f_\lambda t_\lambda / t$. *The element P is represented by the unit matrix I;*

CASE II. $z_\mu = p - 1$, $f_\lambda = 1$, $t_\lambda = t$. *The elements of $\mathfrak{B}$ are represented by scalar multiples of I;*

CASE III. $z_\mu = p - (p-1)/t \geqq (p+1)/2$, $f_\lambda = 1$, $t_\lambda = t \geqq 2$. *The elements of $\mathfrak{B}$ are represented by scalar multiples of I;*

CASE IV. $z_\mu = (p-1) f_\lambda / t$.

The expression S in Theorem 1 *vanishes in all these cases.*

4. A formula for the multiplication of certain characters. Let ω be an irreducible character of $\mathfrak{G}$ such that ω possesses a p-conjugate character ω' with $\omega' \neq \omega$. Then ω will belong to one of the l blocks B_λ of type 0. We must have $t_\lambda \geqq 2$, and ω must be a member of the exceptional family of B_λ. Clearly, we have

$$\omega(G) = \omega'(G) \qquad \text{(for } p\text{-regular elements } G \text{ of } \mathfrak{G}).$$

Consider a character ζ_μ belonging to the same block B_1 as the character $\zeta_1 = 1$. Since $\theta_1 = 1$, $t_1 = t$, the formulae (2) of § 2, applied to ζ_μ, yield

$$\left.\begin{array}{ll} \text{CASES I, II.} & \zeta_\mu(G) = \delta_\mu \\ \text{CASES III, IV.} & \Sigma\zeta_\sigma(G) = \delta_\mu \end{array}\right\} \text{(for } p\text{-singular elements of } \mathfrak{G}),$$

where the sum is extended over all the characters ζ_σ which are p-conjugate to ζ_μ. Accordingly, we have

$$\begin{array}{ll} \text{CASES I, II.} & (\omega(G) - \omega'(G))(\zeta_\mu(G) - \delta_\mu) = 0; \\ \text{CASES III, IV.} & (\omega(G) - \omega'(G))(\Sigma\zeta_\sigma(G) - \delta_\mu) = 0, \end{array}$$

for every G in $\mathfrak{G}$.

In Case I, we have $\delta_\mu = 1$,

$$\omega\zeta_\mu + \omega' = \omega'\zeta_\mu + \omega.$$

This implies $\omega\zeta_\mu \supseteqq \omega$. Consequently, $\omega\bar{\omega} \supseteqq \zeta_\mu$, (cf. Lemma 1). Similarly, in Case II, we have $\delta_\mu = -1$,

$$\omega\zeta_\mu + \omega = \omega'\zeta_\mu + \omega'.$$

Hence $\omega\zeta_\mu \supseteqq \omega'$, $\omega'\bar{\omega} \supseteqq \zeta_\mu$.

In Case III, we have

$$\omega\Sigma\zeta_\sigma + \omega' = \omega'\Sigma\zeta_\sigma + \omega.$$

For at least one ζ_σ belonging to the exceptional family of B_1, $\omega\zeta_\sigma$ contains ω, and hence $\omega\bar{\omega} \supset \zeta_\sigma$.

Similarly, in Case IV, the product $\omega'\bar{\omega}$ contains at least one member of the exceptional family of B_1. This now gives the following result:

THEOREM 2. *Let ω be an irreducible character of $\mathfrak{G}$ which possesses a p-conjugate character $\omega' \neq \omega$. Then we have one of the following two cases*:

CASE $\mathcal{A}$. $\omega\bar{\omega} \supseteqq \chi + \Sigma\alpha, \quad \omega'\bar{\omega} \supseteqq \Sigma\beta;$
CASE $\mathcal{B}$. $\omega\bar{\omega} \supseteqq \Sigma\alpha, \quad \omega'\bar{\omega} \supseteqq \chi + \Sigma\beta.$

Here α ranges over the characters ζ_μ of B_1, for which $\delta_\mu = 1$ and which do not belong to the exceptional family; β ranges over those ζ_μ in B_1, for which $\delta_\mu = -1$, and which do not belong to the exceptional family. Finally, χ is a character ζ_μ, chosen suitably from the exceptional family of B_1; we have CASE *$\mathcal{A}$ if $\delta_\mu = 1$ for this ζ_μ, and* CASE *$\mathcal{B}$ if $\delta_\mu = -1$.*

5. Representations of a degree $n < (p-1)/2$. Preliminary remarks. We now state

THEOREM 3. *Let $\mathfrak{G}$ be a group of order $g = pg'$, with $(p, g') = 1$, which has no normal subgroup of the prime order p. The degree n of any $(1-1)$-representation $\mathfrak{Z}$ of $\mathfrak{G}$ is not smaller than $(p-1)/2$.*

Proof. (a) We first show that it is sufficient to prove the theorem in the case where the representation is irreducible. Let us assume that the theorem is correct for irreducible $\mathfrak{Z}$. Suppose now that $\mathfrak{Z}$ is a $(1-1)$-representation of $\mathfrak{G}$ of a degree $n < (p-1)/2$, which is reducible; let $\mathfrak{F}_1, \mathfrak{F}_2, \cdots, \mathfrak{F}_r$ be the irreducible constituents of $\mathfrak{G}$, and set

$$\mathfrak{Z} = \begin{pmatrix} \mathfrak{F}_1 & & & \\ & \mathfrak{F}_2 & & \\ & & \ddots & \\ & & & \mathfrak{F}_r \end{pmatrix}.$$

Each of the $\mathfrak{F}_\rho$ forms a $(1-1)$-representation of a certain factor group $\mathfrak{U}_\rho$ of $\mathfrak{G}$. Since the degree of $\mathfrak{F}_\rho$ is certainly smaller than $(p-1)/2$, Theorem 3 holds for $\mathfrak{F}_\rho$, if the order of $\mathfrak{U}_\rho$ is divisible by p. Consequently, $\mathfrak{U}_\rho$ contains a normal subgroup of order p. If Y_ρ is a matrix of $\mathfrak{F}_\rho$ representing a generating element of this subgroup, then

$$Y_\rho, Y_\rho^2, \cdots, Y_\rho^{p-1}$$

are the only matrices of order p appearing in $\mathfrak{F}_\rho$. The groups $\mathfrak{F}_\rho$ which represent factor groups $\mathfrak{U}_\rho$ of an order prime to p do not contain matrices of order p; set here $Y_\rho = 1$. Any element of $\mathfrak{G}$ of order p then is represented in $\mathfrak{Z}$ by a matrix

$$X = \begin{pmatrix} Y_1^{h_1} & & & & \\ & Y_2^{h_2} & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & Y_r^{h_r} \end{pmatrix},$$

where the h_ρ are integral rational numbers. Any two matrices X of this type commute. It follows that all the elements of order p in $\mathfrak{G}$ together with the 1-element form a subgroup which is necessarily normal. The order of this subgroup must be a power of p, and since $g \not\equiv 0 \pmod{p^2}$, the order must be p. Hence $\mathfrak{G}$ contains a normal subgroup of order p, as was to be shown.

(b) Suppose that $\mathfrak{Z}$ is an irreducible $(1-1)$ representation of the group $\mathfrak{G}$ of order $g = pg'$, $(g', p) = 1$. We have to show that if the degree n of $\mathfrak{Z}$ is smaller than $(p-1)/2$, then $\mathfrak{G}$ contains a normal subgroup of order p. We may assume that Theorem 3 is true for groups of smaller order than g.

If $\mathfrak{G}$ contains a normal subgroup $\mathfrak{H} \subset \mathfrak{G}$ whose order is divisible by p, then $\mathfrak{Z}$ induces a $(1-1)$-representation of $\mathfrak{H}$. Theorem 3 applied to $\mathfrak{H}$ shows that $\mathfrak{H}$ has a normal subgroup $\mathfrak{P}$ of order p. Since $\mathfrak{P}$ is necessarily a characteristic subgroup of $\mathfrak{H}$, it is a normal subgroup of $\mathfrak{G}$, and the theorem is proven for this case. In particular, the theorem is true for $\mathfrak{G}$, if for the commutator subgroup $\mathfrak{G}'$ of $\mathfrak{G}$, the index $(\mathfrak{G} : \mathfrak{G}')$ is larger than 1, but different from p; under these conditions $\mathfrak{G}$ clearly contains normal subgroups $\mathfrak{H} \subset \mathfrak{G}$ whose order is divisible by p.

(c) Assume next that $(\mathfrak{G} : \mathfrak{G}') = p$. Then $\mathfrak{G}'$ cannot contain an element of order p. It follows from (1) that $P^{\gamma^t} = P$, since $P^{\gamma^t - 1} = P^{-1}M^{-1}PM$ is a commutator element. As γ was a primitive root $\pmod{p}$, we find that $t = p - 1$ and hence $t_\lambda = p - 1$ for $\lambda = 1, 2, \cdots, l$. The block B_λ, which contains the character ω of the representation $\mathfrak{Z}$, consists of the $p-1$ characters which are p-conjugate to ω and one further character ζ_μ; for the latter we have $r_\mu = 1$. Further,[8]

$$(5) \qquad \omega(G) = \zeta_\mu(G) \qquad \text{(for } p\text{-regular elements } G\text{)}.$$

In particular, taking $G = 1$, we see that ζ_μ has the same degree n as $\mathfrak{Z}$. Now, Corollary 2 of § 3 can be applied. Since the character does not belong to the exceptional family of B_λ, we must have Case I, if $n < (p-1)/2$. The representation $\mathfrak{Z}_\mu$ corresponding to ζ_μ represents P by the unit matrix. On the other hand, $\mathfrak{Z}$ is a $(1-1)$ representation, and hence $\omega(G) \neq n$ for $G \neq 1$. For p-regular elements G, (5) yields $\zeta_\mu(G) \neq n$. The elements of $\mathfrak{G}$, which in $\mathfrak{Z}_\mu$ are represented by the unit matrix, form a normal subgroup $\mathfrak{J}$. It now follows that $\mathfrak{J}$ consists of all the conjugates of P^i, $(i = 0, 1, 2, \cdots, p-1)$. Then $\mathfrak{J}$ must have the order p, and Theorem 3 is true for the group $\mathfrak{G}$ in this case.

---

[8] Cf. [4], Theorem 6.

Consequently, Theorem 3 will be proved completely when the following lemma is proved:

Lemma 2. *Let $\mathfrak{G}$ be a group of order $g = pg'$, with $(p, g') = 1$, which is identical with its commutator-subgroup. The degree of any irreducible $(1-1)$-representation of $\mathfrak{G}$ is not smaller than $(p-1)/2$.*

6. Continuation. Proof of Lemma 2. Let $\mathfrak{G}$ be a group of order $g = pg'$ with $(p, g') = 1$. If $\mathfrak{G} = \mathfrak{G}'$, then $\mathfrak{G}$ has no linear character except $\zeta_1 = 1$; we shall use only the fact that the block B_1 contains no linear character except ζ_1. Suppose that $\mathfrak{G}$ has an irreducible $(1-1)$ representation $\mathfrak{Z}$ of degree $n \leqq (p-1)/2$.[9] Theorem 1 and the corollaries can be applied to $\mathfrak{Z}_\mu = \mathfrak{Z}$. For the character ω of $\mathfrak{Z}$, Case IV must hold. We certainly have $t > 1$, and hence $t_\lambda > 1$,

$$t_\lambda \geqq t \geqq 2 \tag{6}$$

If ω belongs to the block B_λ, the value of ω for elements of $\mathfrak{N}$ is

$$\omega = \sum_{\kappa=0}^{\tau_\lambda - 1} \theta_\lambda^{(\kappa)} \sum_{\rho}{}'' (\epsilon)^\rho \qquad (G \text{ in } \mathfrak{N}), \tag{7}$$

where ρ ranges over the q_λ values between 0 and p, whose q_λ-th power is congruent (mod p) to the q_λ-th power of $\gamma^{\kappa t}$, $(q_\lambda = (p-1)/t_\lambda$, $\tau_\lambda = t_\lambda/t)$.

Let ω' be the p-conjugate character obtained by replacing ϵ by ϵ^γ. For elements of $\mathfrak{N}$, we have

$$\omega' = \sum_{\kappa=0}^{\tau_\lambda - 1} \theta_\lambda^{(\kappa)} \sum_{\rho}{}'' (\epsilon)^{\rho\gamma} \qquad (G \text{ in } \mathfrak{N}). \tag{8}$$

No term $(\epsilon)^{\rho\gamma}$ appearing in (8) can also appear in (7), since for the exponents ρ we have $\rho^q \equiv 1$, $(\gamma\rho)^q \equiv \gamma^q \not\equiv 1 \pmod p$, because $q = (p-1)/t < p-1$.

Using the formulae (7) and (8), we find

$$\bar{\omega}\omega' = \sum_{\iota,\kappa} \bar{\theta}_\lambda^{(\iota)} \theta_\lambda^{(\kappa)} \sum_{\rho}{}'' \sum_{\sigma}{}'' (\epsilon)^{\rho\gamma-\sigma} \tag{9}$$

---

[9] We include the equality sign in order to avoid a repetition further on.

for elements G belonging to $\mathfrak{N}$. Expressing $\bar{\theta}_\lambda^{(\iota)}\theta_\lambda^{(\kappa)}$ by the irreducible characters of $\mathfrak{B}$, the product $\bar{\omega}\omega'$ can be written as a linear combination of the irreducible characters of $\mathfrak{N}$. Since $\rho\gamma \not\equiv \sigma \pmod{p}$, every term is of the form $\theta_\lambda^{(j)} \cdot (\epsilon)^\nu$, with $\nu \not\equiv 0 \pmod{p}$.

On the other hand, $\bar{\omega}\omega'$ contains $\Sigma\beta$ in Case $\mathcal{A}$ and $\chi + \Sigma\beta$ in Case $\mathcal{B}$ (for arbitrary elements G of $\mathfrak{G}$). Again, we restrict ourselves to elements G belonging to $\mathfrak{N}$ and express all the characters by the irreducible characters of $\mathfrak{N}$. It follows that every β can contain only terms $\theta_\lambda^{(j)} \cdot (\epsilon)^\nu$ with $\nu \not\equiv 0 \pmod{p}$; in Case $\mathcal{B}$ this will also hold for χ. Consequently, if Theorem 1 is applied to β (and also to χ in Case $\mathcal{B}$), it follows that the expression S vanishes. Since β and χ belong to B_1, and $t_1 = t$, we then have

$$(10) \qquad \beta = \theta_1 \sum_{\rho=1}^{p-1} (\epsilon)^\rho,$$

$$(10, \mathcal{B}) \qquad \chi = \theta_1 \sum (\epsilon)^\rho, \qquad \text{(in Case } \mathcal{B}\text{)},$$

where ρ ranges in $(10, \mathcal{B})$ over the values between 0 and p, for which $\rho^q \equiv 1 \pmod{p}$, $(q = (p-1)/t)$. The character θ_1 is the 1-character of $\mathfrak{B}$; hence β has degree $p-1$, and χ degree $(p-1)/t = q$ in Case $\mathcal{B}$.

The character $\bar{\theta}_\lambda^{(\iota)}\theta_\lambda^{(\kappa)}$ in (9) does not contain θ_1 if $\iota \neq \kappa$; if $\iota = \kappa$, the product $\bar{\theta}_\lambda^{(\iota)}\theta_\lambda^{(\iota)}$ contains θ_1 exactly once. It now follows from (9) that $\omega'\bar{\omega}$ contains exactly $\tau_\lambda q_\lambda^2$ terms $\theta_1 \cdot (\epsilon)^\nu$. Let u be the number of characters β. On comparing (10) and (9), we find

$$(11) \qquad \begin{array}{ll} \text{CASE } \mathcal{A}. & u(p-1) \leqq \tau_\lambda q_\lambda^2 = (p-1)^2/tt_\lambda; \\ \text{CASE } \mathcal{B}. & u(p-1) + (p-1)/t \leqq \tau_\lambda q_\lambda^2 = (p-1)^2/tt_\lambda. \end{array}$$

The total number of characters α and β together is $q = (p-1)/t$, (cf. § 2), and we have, therefore, $q-u$ characters α. One of them, the character $\alpha = \zeta_1 = 1$, is linear. As assumed above, all the others have a degree larger than 1. Corollary 1 then shows that $Dg(\alpha) \geqq p+1$, where $Dg(\Theta)$ denotes the degree of a character Θ. We now treat the cases $\mathcal{A}$ and $\mathcal{B}$ separately, using the same method in each case.

CASE $\mathcal{A}$. We have[10]

$$Dg(\chi) + \sum_\alpha Dg(\alpha) = \sum_\beta Dg(\beta),$$

and $Dg(\chi) \geqq p - (p-1)/t = p - q$, according to Corollary 1, since we have Case III for χ. Consequently

$$(12, \mathcal{A}) \qquad p - q + 1 + (q - u - 1)(p+1) \leqq \Sigma Dg(\beta) = u(p-1),$$

---

[10] Cf. [4], Theorem 6, (49).

and hence $qp \leqq 2up$, $q \leqq 2u$. However, (11) implies $u \leqq (p-1)/tt_\lambda = q/t_\lambda$. Combining these two inequalities and using (6), we obtain

$$(13, \mathfrak{A}) \qquad t_\lambda = 2, \quad t = 2, \quad u = (p-1)/4.$$

In particular, $p \equiv 1 \pmod 4$. Further, in (12) the equality sign must hold, i. e. we have

$$(14, \mathfrak{A}) \qquad Dg(\chi) = (p+1)/2, \quad Dg(\alpha) = p+1 \quad \text{for} \quad \alpha \neq \zeta_1.$$

The number of characters α of degree $p+1$ is

$$q - u - 1 = (p-1)/2 - (p-1)/4 - 1 = (p-5)/4.$$

From (7), we obtain

$$(15) \qquad Dg(\omega) = f_\lambda \tau_\lambda q_\lambda = f_\lambda(p-1)/t = f_\lambda(p-1)/2.$$

This shows that $Dg(\omega) = n < (p-1)/2$ is impossible. If $n = (p-1)/2$, then $f_\lambda = Dg(\theta_\lambda) = 1$.

Case $\mathfrak{B}$: Here

$$\sum_\alpha Dg(\alpha) = Dg(\chi) + \sum_\beta Dg(\beta),$$

$$(12, \mathfrak{B}) \qquad 1 + (q-u-1)(p+1) \leqq Dg(\chi) + \sum Dg(\beta)$$
$$= (p-1)/t + u(p-1) = q + u(p-1).$$

Hence $qp - p \leqq 2up$, $(q-1)/2 \leqq u$, whereas (11) gives $u \leqq q/t_\lambda - 1/t$. Combining the two inequalities, we have $qt_\lambda - t_\lambda \leqq 2q - 2\tau_\lambda$, i. e., $q(t_\lambda - 2) \leqq t_\lambda - 2\tau_\lambda \leqq t_\lambda - 2$. Now $t_\lambda \geqq 2$ by (6) and $q = Dg(\chi) \geqq 2$, since $\zeta_1 = 1$ is the only linear character of B_1. Thus $2(t_\lambda - 2) \leqq t_\lambda - 2$. Again, we find $t_\lambda = 2$, which implies $t = 2$. The two inequalities for u now have the form $(q-1)/2 \leqq u \leqq q/2 - 1/2$. Corresponding to $(13, \mathfrak{A})$ we have here

$$(13, \mathfrak{B}) \qquad t_\lambda = 2, \quad t = 2, \quad u = (p-3)/4.$$

Since u is an integer, this gives $p \equiv 3 \pmod 4$. In $(12, \mathfrak{B})$ the equality sign holds. Consequently,

$$(14, \mathfrak{B}) \qquad Dg(\alpha) = p+1 \quad \text{for} \quad \alpha \neq \zeta_1.$$

The number of characters α of degree $p+1$ is here

$$q - u - 1 = (p-1)/2 - (p-3)/4 - 1 = (p-3)/4.$$

The equation (15) holds as in Case $\mathfrak{A}$. Again, it follows that $Dg(\omega) = n < (p-1)/2$ is impossible. If $n = (p-1)/2$, then $f_\lambda = 1$. This finishes the proof of Lemma 2 and of Theorem 3.

At the same time, we obtain

LEMMA 3. *Let* $\mathfrak{G}$ *be a group of order* $q = pg'$ *with* $(p, g') = 1$ *and assume*[11] *that* ζ_1 *is the only linear character in the block* B_1. *Suppose that* $\mathfrak{G}$ *has an irreducible* $(1-1)$*-representation* $\mathfrak{Z}$ *of degree* $n = (p-1)/2$. *Case* $\mathfrak{A}$: *If* $p \equiv 1 \pmod 4$, *then the first block* B_1 *of characters consists of* $\zeta_1 = 1$, $(p-5)/4$ *characters* α *of degree* $p+1$, *two conjugate characters* χ, χ' *of degree* $(p+1)/2$, *and* $(p-1)/4$ *characters* β *of degree* $p-1$. *Case* $\mathfrak{B}$: *If* $p \equiv 3, (\mathrm{mod}\ 4)$, *then* B_1 *consists of* $\zeta_1 = 1$, $(p-3)/4$ *characters* α *of degree* $p+1$, *two conjugate characters* χ, χ' *of degree* $(p-1)/2$, *and* $(p-3)/4$ *characters* β *of degree* $p-1$. *In either case, we have* $t = 2$, *and* $t_\lambda = 2$, *if the character of* $\mathfrak{Z}$ *belongs to the block* B_λ.

7. The case $n = (p-1)/2$. Preliminary remarks.

We now state

THEOREM 4. *Let* $\mathfrak{G}$ *be a group of order* $g = pg'$ *with* $(p, g') = 1$, *which has no normal subgroup of order* p. *If* $\mathfrak{G}$ *has a* $(1-1)$*-representation* $\mathfrak{Z}$ *of degree* $n = (p-1)/2$, *then the factor group of* $\mathfrak{G}$ *modulo the center* $\mathfrak{C}$ *of* $\mathfrak{G}$ *is isomorphic with* $LF(2, p)$. *In other words:* $\mathfrak{Z}$, *considered as a collineation group, represents* $LF(2, p)$ *isomorphically.*

Proof. (a) If $\mathfrak{Z}$ is reducible, the argument in § 5 (a) can be applied. It follows from Theorem 3 that this case is impossible. We may, therefore, assume that $\mathfrak{Z}$ is irreducible.

(b) We next deal with the case where the block B_1 of characters contains a linear character besides $\zeta_1 = 1$. We first assume that this linear character is one of the characters α. Theorem 1 and Corollary 2 show that $\alpha(G) = 1$ for all p-singular elements of $\mathfrak{G}$. If $\mathfrak{H}$ is the normal subgroup consisting of all elements of $\mathfrak{G}$ with $\alpha(G) = 1$, then $\mathfrak{H}$ contains P and all other elements of order p; further, it contains PV for any V in $\mathfrak{V}$, and hence it contains V, i. e.,

$$\mathfrak{H} \geqq \{P\} \times \mathfrak{V} = \mathfrak{N}. \tag{16}$$

The representation $\mathfrak{Z}$ of $\mathfrak{G}$ induces a representation $\mathfrak{Z}^*$ of degree $(p-1)/2$ of $\mathfrak{H}$. If $\mathfrak{Z}^*$ is reducible, the remark in (a) shows that $\mathfrak{H}$ contains a normal subgroup $\mathfrak{P}$ of order p. According to Sylow's theorem, this subgroup is characteristic, and hence it is a normal subgroup of $\mathfrak{G}$, in contradiction to the assumption.

Suppose now that $\mathfrak{Z}^*$ is irreducible. We may assume that Theorem 4 is true for all groups whose order is smaller than g. Then the theorem holds for $\mathfrak{H}$. The factor group $\mathfrak{H}/\mathfrak{C}_0$ modulo the center $\mathfrak{C}_0$ of $\mathfrak{H}$ is isomorphic with

---

[11] Cf. the remark at the beginning of § 6.

$LF(2, p)$. Since in $LF(2, p)$ every element of order p is congruent to its γ^2-th power (where γ again is a primitive root mod p), we may find an element H in $\mathfrak{H}$ and a center element C_0 such that

$$H^{-1}PH = P^{\gamma^2}C_0.$$

But then $P^{\gamma^2}C_0$ is an element of order p, and since it commutes with P, it must be a power of P. Consequently, C_0 also is a power of P. However, the center $\mathfrak{C}_0$ of $\mathfrak{H}$ cannot contain a subgroup of order p, because any normal subgroup of $\mathfrak{H}$ of order p is a normal subgroup of $\mathfrak{G}$. This shows that C_0 cannot have order p; we have, then, $C_0 = 1$, and P and P^{γ^2} are conjugate in $\mathfrak{H}$. The elements P and P^{γ} are not conjugate in $\mathfrak{G}$, since this would mean $t = 1$, and Theorem 1 and the corollaries show that then $\mathfrak{G}$ cannot have a $(1-1)$-representation of degree $< p - 1$. The elements of order p form two classes of $\mathfrak{H}$, which also are two classes of conjugate elements in $\mathfrak{G}$. For any G in $\mathfrak{G}$, the element $G^{-1}PG$ must be conjugate to P in $\mathfrak{H}$. Then G is congruent to an element of $\mathfrak{H}$ modulo $\mathfrak{N}(P) = \mathfrak{N}$. Because of (16), G belongs to $\mathfrak{H}$, and $\mathfrak{G} = \mathfrak{H}$. Then the character α is the 1-character ζ_1. A linear character $\neq \zeta_1$ of B_1 cannot be one of the characters α.

(c) According to Corollary 2 (§ 3), a linear character of B_1, which is not one of the characters α, must belong to the exceptional family of B_1 and we must have $t = p - 1$. Now the argument in § 5 (c) can be used. It shows that $\mathfrak{G}$ would have a normal subgroup of order p, which contradicts the assumption.

Hence ζ_1 is the only linear character of B_1. We see that $\mathfrak{G}$ satisfies the assumptions of Lemma 3. The degrees of the characters of B_1 will therefore have the values given in Lemma 3. The proof of Theorem 4 is based on a further investigation of these characters of B_1. In § 8, we shall show that all the characters α are real. In § 9, the degrees of the modular characters of B_1 are obtained. One of these degrees is equal to 3. In § 11, the corresponding modular representation is used for the proof of the isomorphism of $\mathfrak{G}/\mathfrak{C}$ and $LF(2, p)$.

(d) We add a few simple remarks. Let ω be the character of the representation $\mathfrak{Z}$ of degree $n = (p-1)/2$. In the notation of § 4, we must have Case IV and (2) takes the form

$$\omega = \theta_\lambda \sum_{\rho} (\epsilon)^{\rho} \quad \text{(for elements of } \mathfrak{N}\text{)},$$

where ρ ranges over the $n = (p-1)/2$ quadratic residues (mod p), and where θ_λ is a linear character of $\mathfrak{B}$. In particular, for elements V of $\mathfrak{B}$ we have $\omega(V) = \theta_\lambda(V)n$, which shows that V is represented by $\theta_\lambda(V)I$ in $\mathfrak{Z}$.

Since $\mathfrak{Z}$ is a $(1-1)$-representation, all the elements V belong to the center $\mathfrak{C}$ of $\mathfrak{G}$. Conversely, any element of $\mathfrak{C}$ belongs to $\mathfrak{N}$. No element of $\mathfrak{C}$ can have order p. Hence $\mathfrak{C} \leqq \mathfrak{B}$, i. e.,

$$\mathfrak{C} = \mathfrak{B}. \tag{17}$$

The representations $\mathfrak{Z} \times \bar{\mathfrak{Z}}$ and $\mathfrak{Z}' \times \bar{\mathfrak{Z}}'$ will represent every V in $\mathfrak{B}$ by the unit matrix, when $\bar{\mathfrak{Z}}$ is the conjugate complex representation, and $\mathfrak{Z}'$ the representation with the p-conjugate character ω'. Then Theorem 2 shows that every representation of the first block B_1 represents the elements of $\mathfrak{B}$ by the unit matrix. Actually, the relations of Theorem 2 are equalities in our case.

$$\begin{aligned} &\text{Case } \mathfrak{A}: \quad \omega\bar{\omega} = \chi + \Sigma\alpha, \quad \omega'\bar{\omega} = \Sigma\beta, \\ &\text{Case } \mathfrak{B}: \quad \omega\bar{\omega} = \Sigma\alpha, \quad \omega'\bar{\omega} = \chi + \Sigma\beta, \end{aligned} \tag{18}$$

since both sides have the same degree on account of Lemma 3 (cf. (12) which turned out to be an equality in our case). If, therefore, an element G of $\mathfrak{G}$ is represented by the unit matrix in all representations of B_1, then (18) shows that $\mathfrak{Z}(G) \times \bar{\mathfrak{Z}}(G) = I$. This implies that $\mathfrak{Z}(G)$ can have only one characteristic root c, i. e., $\mathfrak{Z}(G) = cI$, and G belongs to the center $\mathfrak{B}$ of $\mathfrak{G}$.

On the other hand, every irreducible representation of $\mathfrak{G}/\mathfrak{B}$ defines an irreducible representation of $\mathfrak{G}$. If ζ_ν is the corresponding character, we have $\zeta_\nu(P^jV) = \zeta_\nu(P) \equiv z_\nu \pmod{1-\epsilon}$ for V in $\mathfrak{B}$. We see easily[12] that ζ_ν belongs to B_1, provided that its degree z_ν is prime to p.

8. The characters of the group $\{P, M, \mathfrak{B}\}/\mathfrak{B}$. According to Lemma 3, we have $t = 2$ and (1) thus has the form

$$M^{-1}PM = P^{\gamma^2}.$$

For $n = (p-1)/2$, the n-th power of M belongs to $\mathfrak{B}$. The elements $M, P \pmod{\mathfrak{B}}$ generate a group $\mathfrak{W} = \{P, M, \mathfrak{B}\}/\mathfrak{B}$ of order np. Let ξ be a primitive n-th root of unity. Then

$$M \to \xi, \qquad P \to 1$$

defines a linear representation of $\mathfrak{W}$. If (ξ) is its character, every linear character of $\mathfrak{W}$ is of the form $(\xi)^\nu$, $(\nu = 0, 1, \cdots, n-1)$. Besides, $\mathfrak{W}$ has two conjugate characters[13] X and X' of degree n. These characters vanish for the elements M^κ $(\kappa = 1, 2, \cdots, n-1)$, i. e.,

---

[12] We may apply the condition [1], (27) to ζ_1 and ζ_ν.

[13] This can either be seen directly or as a special case of the results of [4]. For the group $\mathfrak{W}$, we have $l = 1$, $t = 2$. Since we have $n = (p-1)/2$ linear characters, the two p-conjugate characters X, X' must have degree n as $n + 2n^2 = np$ is the order of $\mathfrak{W}$;

$$X(1) = X'(1) = n, \quad X(M^{\kappa}) = X'(M^{\kappa}) = 0 \quad \text{for } \kappa \not\equiv 0 \pmod n.$$

After interchanging X and X', if necessary, we may assume that

$$X = \Sigma(\epsilon)^{\rho} \quad \text{(for the elements of } \{P\}),$$

where ρ ranges over the n quadratic residues $(\bmod\, p)$. In the representation corresponding to X, the element P is represented by a matrix with the characteristic roots ϵ^{ρ}, while M is represented by a matrix whose roots are the n-th roots of unity. The determinant of the former matrix is 1 and that of the latter is $(-1)^{n+1}$.

From Theorem 2, applied to the group $\mathfrak{W}$, it follows that $X\bar{X}$ contains all the characters $(\xi)^{\nu}$. Clearly, $X\bar{X} \supseteqq 2(\xi)^{\nu}$ is impossible, since it would imply that $X(\xi)^{\nu} \supseteqq 2X$ (cf. Lemma 1), and the left side has a smaller degree than the right side.

The representations of $\mathfrak{G}$ of the first block B_1 represent the elements of $\mathfrak{V}$ by the unit matrix (cf. § 7). Consequently, these representations of $\mathfrak{G}$ induce representations of $\mathfrak{W}$. Restricting ourselves to elements of $\{P, M, \mathfrak{V}\}$, the characters $\alpha, \beta, \chi, \chi'$ can be expressed as linear combinations of $X, X', (\xi)^{\nu}$; the coefficients are rational integers $\geqq 0$. On comparing the characteristic roots, we see easily that for elements of $\{P, M, \mathfrak{V}\}$

$$(19) \qquad \beta = X + X'; \quad \alpha = X + X' + (\xi)^{\mu} + (\xi)^{\nu} \text{ for } \alpha \neq \zeta_1,$$

where μ, ν are two of the numbers $0, 1, 2, \cdots, n-1$. The character χ contains one of the characters X, X'. It is equal to this character in Case $\mathcal{B}$ where $Dg(\chi) = Dg(X) = n$, while in Case $\mathcal{A}$ one of the linear characters $(\xi)^{\kappa}$ must be added, since $Dg(\chi) = n + 1 = Dg(X) + 1$.

The representations of $\mathfrak{G}$ belonging to B_1 are all unimodular.[14] Indeed, the determinants of the matrices of such a representation form a linear character Δ of $\mathfrak{G}$, and $\Delta(V) = 1$ for V in $\mathfrak{V}$. Then Δ itself belongs to B_1 (cf. [12]), and hence $\Delta = \zeta_1 = 1$. On comparing the determinants of the matrices representing M, we see that $\mu \equiv -\nu$ in (19) while $\kappa = n/2$ in Case $\mathcal{A}$, i. e.,

$$(20) \qquad \alpha = X + X' + (\xi)^{\mu} + (\xi)^{-\mu};$$

$$(21, \mathcal{A}) \qquad \chi = X + (\xi)^{n/2} \text{ or } \chi = X' + (\xi)^{n/2} \text{ in Case } \mathcal{A};$$

$$(21, \mathcal{B}) \qquad \chi = X \qquad \text{or } \chi = X' \qquad \text{in Case } \mathcal{B},$$

---

no character of type II occurs. For p-regular elements G, the value $X(G)$ is the sum of the n linear characters, which gives $X(M^k) = 0$ for $k \not\equiv 0 \pmod n$, while the value of $X(P)$ is obtained directly from [4], Theorems 4 and 11.

[14] That is to say, all the elements of $\mathfrak{G}$ are represented by matrices whose determinant is 1.

for elements of $\{P, M, \mathfrak{W}\}$. On forming $\chi\bar{\chi}$ for the elements of the same subgroup, we obtain the same linear characters $(\xi)^\nu$ as are contained in $X\bar{X}$ [15] in Case $\mathfrak{B}$, while a further character $(\xi)^0 = 1$ appears in Case $\mathfrak{A}$. Using a remark above, we find

$$(22, \mathfrak{A}) \qquad \chi\bar{\chi} = 2(\xi)^0 + \sum_{\nu=1}^{n-1} (\xi)^\nu + \cdots \quad (\text{Case } \mathfrak{A}),$$

$$(22, \mathfrak{B}) \qquad \chi\bar{\chi} = \sum_{\nu=0}^{n-1} (\xi)^\nu + \cdots \quad (\text{Case } \mathfrak{B}),$$

where the characters, not written down on the right side, are not linear.

On the other hand, Theorem 2 applied to the character χ of $\mathfrak{G}$ shows that $\chi\bar{\chi}$ contains all characters α, and also one of the characters χ, χ' in Case $\mathfrak{A}$. In Case $\mathfrak{A}$, one of the α is $\zeta_1 = 1 = (\xi)^0$. According to Lemma 3 (§ 6), we have $(p-5)/4 = (n/2) - 1$ further characters α of degree $p+1$. The formula (22) now shows that no two of these α can contain the same $(\xi)^\mu$. Furthermore, $\chi\bar{\chi}$, which (as a character of $\mathfrak{G}$) contains χ or χ', contains only one such constituent. In Case $\mathfrak{B}$, we have $(p-3)/4 = (n-1)/2$ characters α of degree $p+1$. Again, $(22, \mathfrak{B})$ shows that no two of them contain the same $(\xi)^\nu$.

Consider α again as a character of $\mathfrak{G}$. The conjugate complex character $\bar{\alpha}$ also belongs to the first block.[16] According to (20), the same linear characters $(\xi)^\mu$ appear in α and $\bar{\alpha}$, if considered as characters of $\mathfrak{W}$. As we have seen, different α never contain the same $(\xi)^\mu$. Hence $\alpha = \bar{\alpha}$. The character $\chi\bar{\chi}$ is real. Since it contains exactly one of the characters χ, χ', this constituent is real. The p-conjugate characters χ, χ' are either both real or both not real. Hence we have

LEMMA 4. *Under the assumptions of Theorem 4, the degrees of the characters of* B_1 *have the values given in Lemma* 3. *All the characters* α *are real. In Case* $\mathfrak{A}$, *the characters* χ, χ' *also are real.*

9. The degrees of the modular characters of B_1. We separate the two cases $\mathfrak{A}$ and $\mathfrak{B}$. The methods applied will be essentially the same in both cases.

CASE $\mathfrak{A}$: The subset B^*_1 of characters ζ_μ of B_1 with $\delta_\mu = 1$ consists of $\zeta_1 = 1$, χ, χ' and $(p-5)/4$ characters α_σ of degree $p+1$ $(\sigma = 1, 2, \cdots, (p-5)/4)$. The block B_1 contains $(p-1)/2$ modular characters [17] ϕ, and each of them appears as a modular constituent in exactly one of the characters ζ_1, χ, α_σ. On the other hand, each ϕ appears as a modular constituent of exactly one of the $(p-1)/4$ characters β of degree $p-1$. This shows that $Dg(\phi) \leqq p-1$. Furthermore, it follows that we have at most $(p-1)/4$ modular characters ϕ in B_1 with $Dg(\phi) > (p-1)/2$. In the modular sense, each of the $(p-5)/4$ characters α_σ must be reducible. The constituents of the α_σ account for at least $(p-5)/2$ modular characters of B_1, the constituents of χ and the character ζ_1 for at least two further modular characters. Since $(p-5)/2 + 2 = (p-1)/2$ is the full number of modular characters of B_1, it follows that every α_σ has exactly two modular constituents, while χ is modular-irreducible. One of the constituents of α_σ has at least the degree $(p+1)/2$, and χ also has the degree $(p+1)/2$. This gives $(p-1)/4$ characters ϕ of a degree $> (p-1)/2$. As a remark above shows, the other $(p-1)/4$ modular characters of B_1 necessarily have a degree $\leqq (p-1)/2$. In particular, every α splits into two modular constituents of different degrees. Since α is real (Lemma 4), and the modular constituents cannot be conjugate complex, both constituents are real. The character χ also is real (Lemma 4). Hence all the modular characters of B_1 are real.

---

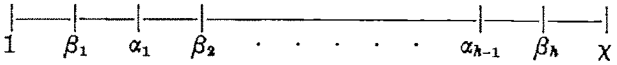
[15] $X\bar{X}$ and $X'\bar{X}'$ contain the same linear characters $(\xi)^\nu$.

[16] This can be seen easily from § 7(d) and also by means of [12].

[17] Cf. [4], Theorem 3.

Now Theorem 14 of [3] can be applied, which shows that the tree corresponding to the block B_1 is an open polygon. The two end points correspond to the modular-irreducible characters 1 and χ. If the notation is chosen suitably, the polygon has the following form $(h = (p-1)/4)$:

(23, $\mathcal{A}$)



The sides of this polygon correspond to the modular characters of B_1. The sum of the degrees of the characters corresponding to adjoining sides is either $p-1$ or $p+1$, according as the common vertex is associated with an α_σ or with a β_σ. The side $1\,\beta_1$ corresponds to the modular character 1, and we find successively, that the degrees of the modular characters of B_1 are given by

$$(24, \mathcal{A}) \quad 1, p-2, 3, p-4, \cdots, (p+5)/2, (p-3)/2, (p+1)/2.$$

Case $\mathcal{B}$: Here, the decomposition of $\zeta_1 = 1$ and the $(p-3)/4$ characters α_σ $(\sigma = 1, 2, \cdots, (p-3)/4)$ of degree $p+1$ into modular constituents gives all the modular constituents ϕ of B_1. Each of these constituents also appears either in χ or in one of the $(p-3)/4$ characters β_σ of degree $p-1$ $(\sigma = 1, 2, \cdots, (p-3)/4)$. This implies $Dg(\phi) \leqq p-1$, and shows also that at most $(p-3)/4$ of the ϕ may have a degree $> (p-1)/2$. Since B_1 contains $(p-1)/2$ modular characters ϕ, it follows as in Case $\mathcal{A}$ that each α_σ splits into two modular-irreducible constituents, one of which has a degree $\geqq (p+1)/2$, while the other must have a degree $\leqq (p-1)/2$.

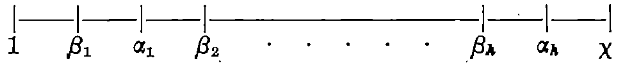
Again, it follows that all the modular characters of B_1 are real. Therefore, the tree corresponding to the block B_1 is an open polygon.

According to Theorem 2, the character $\chi\bar{\chi}$ of $\mathfrak{G}$ contains all the characters α. Since $Dg(\chi) = (p-1)/2$, no other character of $\mathfrak{G}$ can appear (cf. (18, $\mathfrak{B}$)). Hence

$$\chi\bar{\chi} = \zeta_1 + \Sigma\alpha_\sigma.$$

This shows that $\chi\bar{\chi}$ contains the modular 1-character once only. Consequently, χ is modular-irreducible.[18] This shows that $\zeta_1 = 1$ and χ correspond again to the end points of the polygon. Here we have for $p > 3$ $(h = (p-3)/4)$

(23, $\mathfrak{B}$)



The degrees of the modular characters of B_1 are given by

(24, $\mathfrak{B}$) $\quad 1, p-2, 3, p-4, \cdots, (p+3)/2, (p-1)/2.$

The case $p = 3$ is without interest, as the assumptions of Theorem 4 cannot be satisfied. We have therefore the result,

LEMMA 5. *If $\mathfrak{G}$ satisfies the assumptions of Theorem 4, then the first block B_1 contains exactly one modular character of degree 3.*

10. The modular representation of degree 3. Let $\mathfrak{F}$ be the irreducible modular representation of $\mathfrak{G}$ of degree 3, whose character ϕ belongs to the first block B_1 (cf. Lemma 5). We can easily show that $\mathfrak{F}$ can be written with coefficients in the Galois field $GF(p)$ with p elements. Indeed, if this were not so, the traces of the matrices $\mathfrak{F}(G)$, representing G, could not lie in $GF(p)$ for every G in $\mathfrak{G}$.[19] Then we could find a representation $\mathfrak{F}^{\ddagger}$ which is algebraically conjugate to $\mathfrak{F}$ without being similar. It would belong to B_1 and have the degree 3 which contradicts Lemma 5.

Since ϕ was seen to be real, it follows further that $\mathfrak{F}$ is similar to its contragredient representation $\mathfrak{F}^*$. If

$$\mathfrak{F}^{\ddagger} = M^{-1}\mathfrak{F}M,$$

where M is a non-singular matrix with coefficients in $GF(p)$, we conclude easily that MM'^{-1} commutes with every matrix of $\mathfrak{F}$. Hence $MM'^{-1} = cI$ with $c \neq 0$ in $GF(p)$, $M = cM'$, $M' = cM$, and consequently $c^2 = 1$. On the other hand, on forming the determinants of M and M', we find $c^3 = 1$. This gives $c = 1$, i. e., $M' = M$. Consequently, $\mathfrak{F}$ has a non-degenerate quadratic invariant. We may, therefore, assume that the matrices of F are orthogonal matrices of $GF(p)$.

---

[18] The product of two contragredient modular characters contains the 1-character, as in the ordinary theory. Hence if $\chi\bar{\chi}$ contains ζ_1 exactly once in the modular sense, then χ is modular-irreducible. However, if χ is modular-irreducible, $\chi\bar{\chi}$ may contain ζ_1 (modular) more than once.

[19] Cf. footnote 12 of [3].

The group of all orthogonal matrices of degree 3 with coefficients in $GF(p)$ has a normal subgroup $O_1(3, p)$ in Dickson's notation,[20] consisting of the unimodular orthogonal matrices. The factor group is cyclic of order $p - 1$. The elements of $\mathfrak{G}$, which are represented in $\mathfrak{F}$ by matrices belonging to $O_1(3, p)$, form a normal subgroup $\mathfrak{J}$. This group $\mathfrak{J}$ contains $\mathfrak{V}$; the factor group $\mathfrak{G}/\mathfrak{J}$ is cyclic. If $\mathfrak{J} \subset \mathfrak{G}$, we may construct a linear character[21] ζ_μ of $\mathfrak{G}$ such that $\zeta_\mu \neq \zeta_1$ but $\zeta_\mu(V) = 1$ for V in $\mathfrak{V}$. Then ζ_μ would belong to B_1 (§ 7 (d)), while ζ_1 was the only linear character in B_1. Hence all the matrices of $\mathfrak{F}$ belong to $O_1(3, p)$.

The group $O_1(3, p)$ has a normal subgroup $FO(3, p)$ of index 2; we conclude in the same manner that the matrices of $\mathfrak{F}$ belong to $FO(3, p) \cong LF(2, p)$.

Let $\mathfrak{V}^*$ be the normal subgroup of $\mathfrak{G}$ consisting of the elements G which $\mathfrak{F}$ represents by the unit matrix I. Obviously, $\mathfrak{V} \subseteq \mathfrak{V}^*$, and $\mathfrak{G}/\mathfrak{V}^*$ is isomorphic with a subgroup of $LF(2, p)$.

The modular representation $\mathfrak{F}$ appears as a constituent of at least one ordinary irreducible representation $\mathfrak{X}$ of $\mathfrak{G}/\mathfrak{V}^*$. Then $Dg(\mathfrak{X})$ is prime to p, since otherwise $\mathfrak{X}$ would be modular-irreducible,[22] i. e., $Dg(\mathfrak{X}) = Dg(\mathfrak{F}) = 3$ and hence $p = 3$, while we could assume that $p > 3$.

Every representation $\mathfrak{X}$ of $\mathfrak{G}/\mathfrak{V}^*$ yields a representation $\tilde{\mathfrak{X}}$ of $\mathfrak{G}$ consisting of the same matrices as $\mathfrak{X}$ and representing all the elements of $\mathfrak{V}^*$ by the unit matrix. It follows from a remark in § 7 (d) that $\tilde{\mathfrak{X}}$ belongs to the block B_1, since the elements of $\mathfrak{V}$ are represented by I, and since $Dg(\mathfrak{X}) \not\equiv 0 \pmod{p}$. If we let ζ_μ be the character of $\tilde{\mathfrak{X}}$, we certainly have $\zeta_\mu \neq \zeta_1$. The formulae (2) in § 2 show that $\zeta_\mu(P) \neq z_\mu$ which implies that P cannot belong to $\mathfrak{V}^*$. The order v^* of the normal subgroup $\mathfrak{V}^*$ of $\mathfrak{G}$ is not divisible by p. Consequently, the order g/v^* of $\mathfrak{G}/\mathfrak{V}^*$ is divisible by p. Hence the modular representation $\mathfrak{F}$ appears as a modular constituent of two ordinary irreducible representations $\mathfrak{X}_1, \mathfrak{X}_2$ which are not p-conjugate.[23] Each of these two representations $\mathfrak{X}_i$ can be taken for $\mathfrak{X}$ above. The characters of the corresponding representations $\tilde{\mathfrak{T}}_1, \tilde{\mathfrak{T}}_2'$ of $\mathfrak{G}$ are α_1, β_2 for $p > 7$ (cf. (23), (24) $\mathcal{A}, \mathcal{B}$). For $p = 7$, the characters are α_1 and χ; for $p = 5$, they are β_1 and χ. It follows that the order of $\mathfrak{G}/\mathfrak{B}^*$ must be divisible by $(p-1)(p+1)/2$ for $p > 7$, since the group has irreducible representations of degree $p-1$ and $p+1$. The conclusion is also correct for $p = 7$, where we have irreducible representations of degrees 8 and 3, and for $p = 5$, where we have irreducible representations of degrees 4 and 3. Since $\mathfrak{G}/\mathfrak{B}^*$ is isomorphic with a subgroup of $LF(2, p)$, and since its order is divisible by p and by $(p-1)(p+1)/2$, it follows now that

---

[20] L. E. Dickson, *Linear Groups*, Leipzig (1901), Chapter VII.

[21] The modular character corresponding to ζ_μ can be taken as the determinant of $\mathfrak{F}(G)$.

[22] Cf. [1], Theorem 1.

[23] Cf. [2], Theorem 9.

$$\mathfrak{G}/\mathfrak{B}^* \cong LF(2, p). \tag{25}$$

The group $LF(2, p)$ has two classes of conjugate elements which contain elements of order p. Hence[24] $LF(2, p)$ has $(p-1)/2 + 2$ irreducible representations of degrees which are not divisible by p. This is the full number of ordinary representations of $\mathfrak{G}$ belonging to B_1. Consequently, every representation of B_1 represents the elements of $\mathfrak{B}^*$ by the unit matrix. According to § 7 (d), we then have $\mathfrak{B}^* \subseteqq \mathfrak{B}$, and hence

$$\mathfrak{B} = \mathfrak{B}^*. \tag{26}$$

The equations (17), (25), and (26) contain the proof of Theorem 4.

UNIVERSITY OF WISCONSIN.

BIBLIOGRAPHY.

[1] R. Brauer-C. Nesbitt, "On the modular characters of groups," *Annals of Mathematics*, vol. 42 (1941), pp. 556-590.

[2] R. Brauer, "On the connection between the ordinary and the modular characters of groups of finite order," *Annals of Mathematics*, vol. 42 (1941), pp. 926-935.

[3] R. Brauer, "Investigations on group characters," *Annals of Mathematics*, vol. 42 (1941), pp. 936-958.

[4] R. Brauer, "On groups whose order contains a prime number to the first power I," *American Journal of Mathematics*, vol. 64 (1942), pp. 401-420.

---

[24] Actually, the characters of $LF(2, p)$ are known, but the fact used here follows at once from the results of [4] and does not require the knowledge of the characters of $LF(2, p)$.

INDEFINITELY DIFFERENTIABLE FUNCTIONS IN SEVERAL REAL VARIABLES.*

IRWIN E. PERLIN.

1. Introduction.

Let us consider a real single-valued function $F(x)$ of a single real variable x, defined on the interval $a \leq x \leq b$. We assume further that $F(x)$ is indefinitely differentiable on that interval. Let $M_k = \text{l.u.b} \mid F^{(k)}(x) \mid$ on $a \leq x \leq b$.[1] In 1912 Hardy and Littlewood[2] initiated the problem of finding the relationship existing between the positive numbers M_k. Landau,[3] Hadamard,[4] Neder,[5] and Carleman[6] contributed to this problem. Recently Ore,[7] Kolmogoroff,[8] Gorny,[9] Perlin,[10] and Cartan[11] published papers on this problem. Kolmogoroff solved completely for the infinite interval $-\infty < x < \infty$ the problem initiated by Hardy and Littlewood. He proved

---

* Received May 23, 1941; Presented to the American Mathematical Society, December 30, 1940.

[1] By $F^{(0)}(x)$ we mean $F(x)$.

[2] G. H. Hardy and J. E. Littlewood, "Contributions to the arithmetic theory of series," *Proceedings of the London Mathematical Society*, Ser. 2, vol. 11 (1912), pp. 411-478.

[3] E. Landau, "Einige Ungleichungen für zweimal differentiierbare Funktionen," *Proceedings of the London Mathematical Society*, Ser. 2, vol. 13 (1913), pp. 43-49.

[4] J. Hadamard, "Sur le module maximum d'une fonction et de ses dérivées, *Comptes Rendus des Séances de la Societé Mathématique de France* (1914), pp. 68-72.

[5] L. Neder, "Abschätzungen für die Ableitungen einer reellen Funktion eines reellen Arguments," *Mathematische Zeitschrift*, vol. 31 (1930), pp. 356-365.

[6] T. Carleman, *Les fonctions quasi-analytiques*, Paris (1926).

[7] O. Ore, "On functions with bounded derivatives," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 321-326.

[8] A. Kolmogoroff, "Une généralisation de l'inégalité de M. J. Hadamard entre les bornes supérieures des dérivées successives d'une fonction," *Comptes Rendus de l'Académie des Sciences de Paris*, T. 207 (1938), p. 764.

[9] A. Gorny, "Contribution à l'étude des fonctions dérivables d'une variable réelle," *Acta Mathematica*, vol. 71 (1939), pp. 317-358.

[10] I. E. Perlin, "Indefinitely differentiable functions with prescribed least upper bounds," *Bulletin of the American Mathematical Society*, vol. 46 (1940), pp. 272-273.

[11] H. Cartan, "Sur les inégalités entre les maxima des dérivées successives d'une fonction," *Comptes Rendus de l'Académie des Sciences de Paris*, T. 210 (1940), pp. 431-434.

that a necessary and sufficient condition that the triple of positive numbers M_0, M_k, and M_n $(0 < k < n)$ be the least upper bounds of $|F(x)|$, $|F^{(k)}(x)|$, and $|F^{(n)}(x)|$, respectively, of an indefinitely differentiable function $F(x)$ is that

$$M_k^n \leq C_{n,k}^n M_0^{n-k} M_n^k,$$

where

$$C_{n,k} = \frac{T_{n-k}}{T_n^{(n-k)/k}}, \qquad T_n = \frac{4}{\pi} \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)^{n+1}} \qquad (n \text{ even}),$$

$$T_n = \frac{4}{\pi} \sum_{j=0}^{\infty} \frac{1}{(2j+1)^{n+1}} \qquad (n \text{ odd}).$$

For a finite interval Gorny[12] proved that

$$M_k < 4e^{2k}(n/k)^k M_0^{1-(k/n)} M'^{k/n}_n, \qquad M'_n = \max(M_n, M_0 n!\, \delta^{-n}),$$

where δ is the length of the interval.

In this paper we consider real single-valued indefinitely differentiable functions $F(x_1, x_2, \cdots, x_n)$ of several real variables $(x_1, x_2, \cdots, x_n)$ defined on $0 \leq x_p \leq a_p$; $p = 1, 2, \cdots, n$. Let

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b.} \left| \frac{\partial^{r_1+r_2+\cdots+r_n} F(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

on $0 \leq x_p \leq a_p$; $p = 1, 2, \cdots, n$. We shall establish sufficient conditions that there exist an indefinitely differentiable function $F(x_1, x_2, \cdots, x_n)$ of the n real variables $(x_1, x_2, \cdots, x_n)$ defined on $0 \leq x_p \leq a_p$; $p = 1, 2, \cdots, n$, such that the positive numbers $M_{r_1, r_2, \ldots, r_n}$ be the least upper bounds of

$$\left| \frac{\partial^{r_1+r_2+\cdots+r_n} F(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

on $0 \leq x_p \leq a_p$; $p = 1, 2, \cdots, n$ for $r_1, r_2, \cdots, r_n = 0, 1, \cdots \infty$.

2. Polynomials in several real variables. In a previous paper[13] we showed that if $\{M_k\}$ is a given sequence of positive numbers such that the sequence $\{M_k a^k\}$ is monotonically decreasing, then there exists a real single-valued indefinitely differentiable function $F(x)$ of the single real variable x defined on $0 \leq x \leq a$ such that

---

[12] A. Gorny, *loc. cit.*, p. 326.

[13] I. E. Perlin, *loc. cit.*, p. 273.

$$M_k = \text{l. u. b. } | F^{(k)}(x) |$$

on $0 \leq x \leq a$. We wish to state here a condition equivalent to the preceeding. Let

$$S_{r;n} = \sum_{j=0}^{n-r} (-1)^j \frac{M_{j+r}}{j!} a^j .$$

Then the condition that the sequence $\{M_k a^k\}$ be monotonically decreasing is obviously equivalent to the condition that $S_{r;n} > 0$, $r = 0, 1, \cdots, n$; $n = 0, 1, \cdots \infty$. It is this latter condition that we are able to generalize to secure sufficient conditions for the existence of a real single valued indefinitely differentiable function of several real variables with prescribed least upper bounds for the various partial derivatives.

Let $M_{r_1, r_2, \ldots, r_n}$; $r_p = 0, 1, \cdots, m_p$; $p = 1, 2, \cdots, n$ be a set of positive constants. Let

$$S_{r_1, r_2, \ldots, r_n;\, m_1, m_2, \ldots, m_n} = \sum_{i=0}^{m_1-r_1} \sum_{j=0}^{m_2-r_2} \cdots \sum_{k=0}^{m_n-r_n} (-1)^{i+j+\cdots+k} \frac{M_{i+r_1, j+r_2, \ldots, k+r_n}}{i!\, j! \cdots k!} a_1^i a_2^j \cdots a_n^k .$$

We can establish the following theorem.

THEOREM 1. *If* $S_{r_1, r_2, \ldots, r_n;\, m_1, m_2, \ldots, m_n} > 0$, $r_p = 0, 1, \cdots, m_p$; $p = 1, 2, \cdots, n$, *then there exists a polynomial* $P(x_1, x_2, \cdots, x_n)$ *of degree* m_p *in* x_p, $p = 1, 2, \cdots, n$ *such that*

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b. } \left| \frac{\partial^{r_1+r_2+\cdots+r_n} P(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right| ,$$

$0 \leq x_p \leq a_p$; $r_p = 0, 1, \cdots, m_p$; $p = 1, 2, \cdots, n$.

Let us consider the polynomial

$$P(x_1, x_2, \cdots, x_n) = \sum_{r_1=0}^{m_1} \sum_{r_2=0}^{m_2} \cdots \sum_{r_n=0}^{m_n} \frac{S_{r_1, r_2, \ldots, r_n; m_1, m_2, \ldots, m_n}}{r_1!\, r_2! \cdots r_n!} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}.$$

Now,

$$P(a_1, a_2, \cdots, a_n) = \sum_{r_1=0}^{m_1} \sum_{r_2=0}^{m_2} \cdots \sum_{r_n=0}^{m_n} \frac{S_{r_1, r_2, \ldots, r_n; m_1, m_2, \ldots, m_n}}{r_1!\, r_2! \cdots r_n!} a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n},$$

$$= \sum_{r_1=0}^{m_1} \sum_{r_2=0}^{m_2} \cdots \sum_{r_n=0}^{m_n} \frac{1}{r_1!\, r_2! \cdots r_n!} \left(\sum_{i=0}^{m_1-r_1} \sum_{j=0}^{m_2-r_2} \cdots \sum_{k=0}^{m_n-r_n} (-1)^{i+j+\ldots+k} \frac{M_{i+r_1, j+r_2, \ldots, k+r_n} a_1^{i} a_2^{j} \cdots a_n^{k}}{i!\, j! \cdots k!} \right) a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n},$$

$$= \sum_{r_1=0}^{m_1} \sum_{r_2=0}^{m_2} \cdots \sum_{r_n=0}^{m_n} \left(\sum_{i=0}^{m_1-r_1} \sum_{j=0}^{m_2-r_2} \cdots \sum_{k=0}^{m_n-r_n} \frac{(-1)^{i+j+\ldots+k}}{i!\, j! \cdots k!\, r_1!\, r_2! \cdots r_n!} \right) M_{i+r_1, j+r_2, \ldots, k+r_n} a_1^{i+r_1} a_2^{j+r_2} \cdots a_n^{k+r_n},$$

$$= \sum_{t_1=0}^{m_1} \sum_{t_2=0}^{m_2} \cdots \sum_{t_n=0}^{m_n} \left(\sum_{s_1=0}^{t_1} \sum_{s_2=0}^{t_2} \cdots \sum_{s_n=0}^{t_n} \frac{(-1)^{s_1+s_2+\ldots s_n}}{s_1!(t_1-s_1)!\, s_2!(t_2-s_2)! \cdots s_n!(t_n-s_n)!} \right) M_{t_1, t_2, \ldots, t_n} a_1^{t_1} a_2^{t_2} \cdots a_n^{t_n},$$

$$= M_{0,0,\ldots,0}.$$

Further,

$$\frac{\partial^{u_1+u_2+\ldots+u_n} P(x_1, x_2, \cdots x_n)}{\partial x_1^{u_1} \partial x_2^{u_2} \cdots \partial x_n^{u_n}} = \sum_{r_1=u_1}^{m_1} \sum_{r_2=u_2}^{m_2} \cdots \sum_{r_n=u_n}^{m_n} \frac{S_{r_1, r_2, \ldots, r_n; m_1, m_2, \ldots, m_n}}{(r_1-u_1)!(r_2-u_2)! \cdots (r_n-u_n)!} x_1^{r_1-u_1} x_2^{r_2-u_2} x_n^{r_n-u_n}.$$

And similarly,

$$\frac{\partial^{u_1+u_2+\ldots+u_n} P(a_1, a_2, \cdots, a_n)}{\partial x_1^{u_1} \partial x_2^{u_2} \cdots \partial x_n^{u_n}}$$

$$= \sum_{t_1=0}^{m_1-u_1} \sum_{t_2=0}^{m_2-u_2} \cdots \sum_{t_n=0}^{m_n-u_n} \left(\sum_{s_1=0}^{t_1} \sum_{s_2=0}^{t_2} \cdots \sum_{s_n=0}^{t_n} \frac{(-1)^{s_1+s_2+\ldots+s_n}}{s_1!(t_1-s_1)!\, s_2!(t_2-s_2)! \cdots s_n!(t_n-s_n)!} \right) M_{t_1+u_1, t_2+u_2, \ldots, t_n+u_n} a_1^{t_1} a_2^{t_2} \cdots a_n^{t_n},$$

$$= M_{u_1, u_2, \ldots, u_n}$$

Since $P(x_1, x_2, \cdots, x_n) > 0 \; 0 \leq x_p \leq a_p;\; p = 1, 2, \cdots, n$, and

$$\frac{\partial^{u_1+u_2+\cdots+u_n} P(x_1, x_2, \cdots, x_n)}{\partial x_1^{u_1} \partial x_2^{u_2} \cdots \partial x_n^{u_n}} > 0$$

$0 \leq x_p \leq a_p;\; u_p = 0, 1, \cdots, m_p;\; p = 1, 2, \cdots, n$, it follows that

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b.} \left| \frac{\partial^{r_1+r_2+\cdots+r_n} P(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

$0 \leq x_p \leq a_p;\; r_p = 0, 1, \cdots, m_p;\; p = 1, 2, \cdots, n$.

3. Indefinitely differentiable functions with prescribed least upper bounds. In this section we shall establish sufficient conditions that there exist indefinitely differentiable function of several variables with prescribed least upper bounds. Let $M_{r_1, r_2, \ldots, r_n}$, $r_p = 0, 1, \cdots \infty;\; p = 1, 2, \cdots, n$ be an infinite set of positive constants. Let

$$S_{r_1, r_2, \ldots, r_n;\, m_1, m_2, \ldots, m_n} = \sum_{i=0}^{m_1-r_1} \sum_{j=0}^{m_2-r_2} \cdots \sum_{k=0}^{m_n-r_n} (-1)^{i+j+\cdots+k} \frac{M_{i+r_1, j+r_2, \ldots, k+r_n}}{i!\, j! \cdots k!} a_1^{i} a_2^{j} \cdots a_n^{k}$$

$r_p = 0, 1, \cdots, m_p;\quad m_p = 0, 1, \cdots \infty;\quad p = 1, 2, \cdots, n$. We prove the following theorem.

THEOREM 2. *If* $S_{r_1, r_2, \ldots, r_n;\, m_1, m_2, \ldots, m_n} > 0$, $r_p = 0, 1, \cdots, m_p;\; m_p = 0, 1, \cdots, \infty;\; p = 1, 2, \cdots, n$, *then there exists an indefinitely differentiable function* $F(x_1, x_2, \cdots, x_n)$ *such that*

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b.} \left| \frac{\partial^{r_1+r_2+\cdots+r_n} F(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

$0 \leq x_p \leq a_p;\; r_p = 0, 1, \cdots \infty;\; p = 1, 2, \cdots, n$.

By Theorem 1 there exists a sequence $\{P_m(x_1, x_2, \cdots, x_n)\}$ of polynomials in the variables $(x_1, x_2, \cdots, x_n)$ of degree m in each variable such that

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b.} \left| \frac{\partial^{r_1+r_2+\cdots+r_n} P_m(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

$0 \leq x_p \leq a_p;\quad r_p = 0, 1, \cdots, m;\quad p = 1, 2, \cdots, n$. Since the sequences

$$\left\{ \frac{\partial P_m(x_1, x_2, \cdots, x_n)}{\partial x_1} \right\}, \quad \left\{ \frac{\partial P_m(x_1, x_2, \cdots, x_n)}{\partial x_2} \right\}, \cdots, \left\{ \frac{\partial P_m(x_1, x_2, \cdots, x_n)}{\partial x_n} \right\}$$

are uniformly bounded on $0 \leq x_p \leq a_p;\; p = 1, 2, \cdots, n$, the sequence $\{P_m(x_1, x_2, \cdots, x_n)\}$ is equi-continuous on $0 \leq x_p \leq a_p;\; p = 1, 2, \cdots, n$.

Hence, there exists a subsequence $\{P_{m_k}{}^{(1)}(x_1, x_2, \cdots, x_n)\}$ which converges uniformly to a function $F(x_1, x_2, \cdots, x_n)$.[14] Since the sequences

$$\left\{ \frac{\partial^{u_1+u_2+\cdots+u_n} P_m(x_1, x_2, \cdots, x_n)}{\partial x_1^{u_1} \partial x_2^{u_2} \cdots \partial x_n^{u_n}} \right\}$$

are each uniformly bounded on $0 \leq x_p \leq a_p$; $u_p = 0, 1, \cdots, m$; $m = 0, 1, \cdots \infty$; $p = 1, 2, \cdots, n$, we can select a sequence of subsequences

$$\{P_{m_k}{}^{(j)}(x_1, x_2, \cdots, x_n)\}; \qquad j = 1, 2, \cdots \infty$$

such that $\{P_{m_k}{}^{(j)}(x_1, x_2, \cdots, x_n)\}$ is a subsequence of $\{P_{m_k}{}^{(j-1)}(x_1, x_2, \cdots, x_n)\}$, $\{P_{m_k}{}^{(j)}(x_1, x_2, \cdots, x_n)\}$ converges uniformly to $F(x_1, x_2, \cdots, x_n)$, and the sequences of the various partial derivatives of order $\leq j - 1$ of $\{P_{m_k}{}^{(j)}(x_1, x_2, \cdots, x_n)\}$ converge uniformly to the corresponding partial derivatives of $F(x_1, x_2, \cdots, x_n)$. By selecting the diagonal sequence

$$\{P_{m_j}{}^{(j)}(x_1, x_2, \cdots, x_n)\}; \qquad j = 1, 2, \cdots \infty$$

we obtain a sequence of polynomials converging uniformly to $F(x_1, x_2, \cdots, x_n)$ and such that the sequences of the partial derivatives of all orders of this sequence converge uniformly to the corresponding partial derivatives of $F(x_1, x_2, \cdots, x_n)$. It follows then that

$$M_{r_1, r_2, \ldots, r_n} = \text{l. u. b.} \left| \frac{\partial^{r_1+r_2+\cdots+r_n} F(x_1, x_2, \cdots, x_n)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right|$$

$0 \leq x_p \leq a_p$; $r_p = 0, 1, \cdots \infty$; $p = 1, 2, \cdots, n$.

ILLINOIS INSTITUTE OF TECHNOLOGY.

---

[14] E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 2, Cambridge (1926), p. 168.

A PROBLEM IN HEAT CONDUCTION AND AN EXPANSION THEOREM.*

By R. E. Gaskell.

---

1. **Introduction.** Consider a uniform, isotropic, right-cylindrical bar with plane faces, whose lateral surface is thermally insulated. Suppose that this bar has an initial distribution of temperature depending only on the longitudinal coordinate, x. Let the ends of the bar be placed in contact with fixed quantities of well-stirred liquids which are in contact with external media which are maintained at constant temperatures. The problem is: given the initial temperatures of the bar and of the liquids, with the physical constants involved, to find the temperatures of points in the bar or of the liquid at any subsequent time.

Some cases of this problem have been treated by Langer [6],[1] Lowan [7], March and Weaver [8], Peddie [9], Peek [10], and Schumann [11]. In this paper the theory of the Laplace transformation will be used not only to obtain the solution of the problem, but also to establish that solution rigorously and to examine its uniqueness.

Let the faces of the cylindrical bar be at $x = -1$, $x = 1$; and suppose the unit of time so chosen that the diffusivity is unity. The boundary value problem for the temperature function, $U(x,t)$, is the following:

$$(1.1)\qquad U_{xx}(x,t) = U_t(x,t) \qquad (t > 0, -1 < x < 1)$$

$$(1.2)\qquad U_x(-1+0,t) = K_1U(-1+0,t) + C_1U_t(-1+0,t) \qquad (t > 0)$$

$$(1.3)\qquad -U_x(1-0,t) = K_2U(1-0,t) + C_2U_t(1-0,t) \qquad (t > 0)$$

$$(1.4)\qquad U(x,+0) = f(x) \qquad (-1 < x < 1)$$

$$(1.5)\qquad \lim_{t\to 0} U(-1+0,t) = S_1$$

$$(1.6)\qquad \lim_{t\to 0} U(1-0,t) = S_2.$$

In this statement of the problem, $C_1 = M_1q_1/KA$, and $K_1 = k_1B_1/KA$; where, for the liquid on the left, M_1 is the mass, q_1 the specific heat, S_1 the initial temperature, B_1 the area of liquid exposed to the external medium on the left, and k_1 is the external conductivity constant $(k_1 \geqq 0)$; and where K is the

---

* Received February 19, 1941; Revised July 24, 1941; Presented to the American Mathematical Society, April 12, 1940.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

conductivity of the bar and A its cross-sectional area. The constants C_2, K_2 and S_2 are defined analogously for the right end of the bar.[2]

Both C_1 and C_2 are either positive or zero. If both are zero, the problem can be solved by the classical method of using the orthogonal functions generated by a Sturm-Liouville system ([5], Chapters 10, 11). Hence, in the rest of the discussion, let us assume that $C_1 > 0$ or $C_2 > 0$.

2. Application of the Laplace transformation.

The Laplace transformation, applied to (1. 1)-(1. 6), leads to the system

$$\begin{gathered} u_{xx}(x,s) - su(x,s) + f(x) = 0, \\ u_x(-1,s) = K_1 u(-1,s) + C_1 s u(-1,s) - C_1 S_1, \\ -u_x(1,s) = K_2 u(1,s) + C_2 s u(1,s) - C_2 S_2; \end{gathered}$$

$u(x,s)$ being the Laplace transform of $U(x,t)$; i. e.,

$$u(x,s) = \int_0^\infty e^{-st} U(x,t)\,dt.$$

The solution of this system is found to be

$$\begin{aligned} (2.1) \qquad u(x,s) = \{ g_2(1-x,s)[C_1S_1 + \int_{-1}^{1} g_1(1+\xi,s) f((\xi)\,d\xi/\sqrt{s}] \\ + C_2S_2g_1(1+x,s)\}/G(s) + \nu(x,s), \end{aligned}$$

where

$$g_i(x,s) = (C_i s + K_i)\sinh x\sqrt{s} + \sqrt{s}\cosh x\sqrt{s} \qquad (i = 1, 2)$$

$$h_i(x,s) = (C_i s + K_i)\cosh x\sqrt{s} + \sqrt{s}\sinh x\sqrt{s} \qquad (i = 1, 2)$$

$$\nu(x,s) = \left[\int_x^1 \sinh(x-\xi)\sqrt{s}\, f(\xi)\,d\xi\right]/\sqrt{s},$$

$$G(s) = h_1(1,s)g_2(1,s) + g_1(1,s)h_2(1,s).$$

The inverse Laplace transformation cannot readily be applied to the solution (2. 1), so an alternative expression will be found. Let us assume that $f(x)$ is piecewise continuous when $-1 \leqq x \leqq 1$, and that $f'(x)$ is bounded and integrable. For convenience, take $x = a$ as the only point of discontinuity of $f(x)$, and let $J_x = f(x+0) - f(x-0)$, so that $J_x = 0$ when $x \neq a$. After integration by parts, we arrive at another expression for $u(x,s)$ which lends itself more readily to inversion,

---

[2] If the temperatures of the external media are not zero, addition of the proper linear function of x to $U(x,t)$ will give a function, $V(x,t)$, which will satisfy conditions (1. 1)-(1. 6).

$$(2.2)\qquad u(x,s)-\bar{f}(x)/s=[1/sG(s)]\{J(x,s)+I(x,s)$$
$$+[h_2(-x,s)\sinh\sqrt{s}+g_2(-x,s)\cosh\sqrt{s}][C_1S_1s-f(-1)(C_1s+K_1)]$$
$$+[h_1(x,s)\sinh\sqrt{s}+g_1(x,s)\cosh\sqrt{s}][C_2S_2s-f(1)(C_2s+K_2)]\},$$

where

$$I(x,s)=\tfrac{1}{2}\left\{\int_{-1}^{1}g_1(x,s)h_2(-\xi,s)-h_1(x,s)g_2(\xi,s)]f'(\xi)\,d\xi\right.$$
$$+\int_{x}^{1}[g_1(1,s)h_2(1+x-\xi,s)+g_2(1+x-\xi,s)h_1(1,s)]f'(\xi)\,d\xi$$
$$\left.-\int_{-1}^{x}[g_1(1,s)h_2(1-x+\xi,s)+g_2(1-x+\xi,s)h_1(1,s)]f'(\xi)\,d\xi\right\};$$

$$\bar{f}(x)=\tfrac{1}{2}[f(x+0)+f(x-0)],$$

and

$$J(x,s)=\pm\tfrac{1}{2}J_a[h_1(1,s)g_2(1-|x-a|,s)+g_1(1,s)h_2(1-|x-a|,s)]$$
$$-\tfrac{1}{2}J_a[h_1(1,s)g_2(1-x-a,s)-g_1(1,s)h_2(1-x-a,s)]+\tfrac{1}{2}J_xG(s).$$

In the definition of $J(x,s)$, the upper sign is to be used when $x<a$, the lower sign when $x\geqq a$.

3. The solution in integral form. We shall first make several observations concerning the order properties[3] of $[u(x,s)-\bar{f}(x)/s]$. Writing $s=re^{i\theta}$, we see, for example, that when $-\theta_0\leqq\theta\leqq\theta_0<\pi$,

$$[\cosh x\sqrt{s}\cosh y\sqrt{s}]/G(s)=O(s^{-2}\exp\{(|x|+|y|-2)\sqrt{s}\}).$$

Because the integrands occurring in $I(x,s)$ contain such functions, it is easy to see how order properties can be developed for them. In particular, for $-1\leqq\xi\leqq x\leqq 1$,

$$[g_1(1,s)h_2(1-x+\xi,s)+g_2(1-x+\xi,s)h_1(1,s)]/G(s)$$
$$=O(\exp\{-(x-\xi)\sqrt{r}\cos\theta/2\})\quad\text{when}\quad 0\leqq(x-\xi)\leqq 1;$$
$$=O(\exp\{x-\xi-2)\sqrt{r}\cos\theta/2\})\quad\text{when}\quad 1\leqq(x-\xi)\leqq 2.$$

Using this result, we can find the order of the last of the integrals occurring in $I(x,s)$; and eventually this process shows that $I(x,s)/[sG(s)]=O(s^{-3/2})$ when $-\theta_0\leqq\theta\leqq\theta_0$, $-1\leqq x\leqq 1$.

For any fixed value of x such that $-1<x<1$, it can be seen that

---

[3] If there exist positive constants, M and R, such that, for every s for which $|s|\geqq R$, $|s^kF(s)|\leqq M$, then $F(s)$ is of order (s^{-k}), written $F(s)=O(s^{-k})$.

$J(x,s)/[sG(s)] = O(s^{-3/2})$. Since all other terms of the right member of (2.2) are of order $(s^{-3/2})$,

$$u(x,s) - \bar{f}(x)/s = O(s^{-3/2}) \qquad (-1 < x < 1, -\theta_0 \leqq \theta \leqq \theta_0).$$

With this information we can make use of a known theorem ([3], p. 612, Theorem 10), and write the inverse of $[u(x,s) - f(x)/s]$ in integral form. If we denote $L^{-1}\{u(x,s)\}$ by $U(x,t)$, then[4]

$$U(x,t) = \bar{f}(x) + L_i^{-1}\{u(x,s) - \bar{f}(x)/s\}. \tag{3.1}$$

4. A general expansion theorem. In Article 5 we shall develop an expansion of $L_i^{-1}\{u(x,s) - \bar{f}(x)/s\}$ in series by using the following expansion theorem:[5]

THEOREM A. *Let* $f(s)$ *be analytic in each finite region of the complex plane except for simple real poles,* s_j $(j = 0, 1, 2, \cdots)$ *so distributed that a real constant* P *exists such that* $P > s_0 > s_1 > \cdots > s_n > \cdots$. *Also, for some fixed integer* n_0 *and positive constants* M, r_0, A, *let* $f(s)$ *satisfy the conditions:*

(a) *for each* $\theta_1 < \pi$ *a constant* $M_1(\theta_1)$ *exists for which* $|s^k f(s)| < M_1$ *when* $-\theta_1 \leqq \theta \leqq \theta_1$ *and* $|s| \geqq r_0 \qquad (k > 1)$;

(b) $|sf(s)| < M$ *when* s *is on any of the parabolas*

$$r = H_n \csc^2 \theta/2 \qquad (n = n_0, n_0 + 1, \cdots), \tag{4.1}$$

where the real constants H_n *are determined so that* $s_{n+1} < -H_n < s_n$.

Then the inverse transform of $f(s)$ *is given by*

$$F(t) = \sum_{j=0}^{\infty} A_j e^{s_j t} \qquad (t \geqq 0), \tag{4.2}$$

where A_j *is the residue of* $f(s)$ *at* s_j; *and* $F(0) = F(+0) = 0$.

Proof. Under the preliminary conditions and conditions (a) it has been shown ([3], p. 612, Theorem 10) that the inversion integral along the line $R(s) = P$ converges for $t \geqq 0$ to the inverse transform of $f(s)$, that $F(t)$ is continuous for each $t \geqq 0$, and that $F(0) = 0$. Furthermore, under the additional conditions (b), it is known ([4], p. 745, Theorem 1) that for each $t > 0$,

---

[4] $L_i^{-1}\{\phi(s)\}$ is an integral form of the inverse of the Laplace transformation; $L_i^{-1}\{\phi(s)\} = (2\pi i)^{-1} \int_{q-i\infty}^{q+i\infty} e^{tz}\phi(z)\,dz$

[5] Due to Prof. R. V. Churchill.

$$F(t) = L_t^{-1}\{f(s)\} = \sum_{j=0}^{\infty} A_j e^{s_j t}.$$

So the theorem will be proved if it is shown that this series converges to the inversion integral at $t = 0$.

Since the poles at s_j are the only singularities of the function $f(s)$,

$$\left[\int_{P-iQ_n}^{P+iQ_n} f(s)\,ds + \int_{\lambda_n} f(s)\,ds\right] / 2\pi i = \sum_{j=0}^{n} A_j, \tag{4.3}$$

where λ_n is the arc of a parabola (4.1) to the left of the line $R(s) = P$, and $P + iQ_n$, $P - iQ_n$ are the points of intersection of the same parabola with the line $R(s) = P$. We can take n_0 large enough so that $|\theta| > \pi/4$ on the arcs λ_n, and so

$$\begin{aligned} |I_n| = \left|\int_{\lambda_n} f(s)\,ds\right| &\leqq \int_{\pi/4}^{\theta_1} M_1 |s^{-k}|\,|ds| + \int_{\theta_1}^{\pi} M |s^{-1}|\,|ds| \\ &+ \int_{-\pi}^{-\theta_1} M |s^{-1}|\,|ds| + \int_{-\theta_1}^{-\pi/4} M_1 |s^{-k}|\,|ds|. \end{aligned}$$

On λ_n, $|ds| = r\,|\csc\theta/2|\,d\theta$; hence,

$$|I_n| \leqq 2\int_{\pi/4}^{\theta_1} M_1 H_n^{1-k}\,d\theta + 2\int_{\theta_1}^{\pi} M \csc\theta/2\,d\theta.$$

It follows that for any $\epsilon > 0$, $(\pi - \theta_1)$ and $1/n$ can be chosen so small that $|I_n| \leqq \epsilon$. Returning to (4.3), we have

$$\lim_{n\to\infty}\left[\int_{P-iQ_n}^{P+iQ_n} f(s)\,ds\right] / 2\pi i = \lim_{n\to\infty}\sum_{j=0}^{n} A_j. \tag{4.4}$$

Since the infinite integral here is $[L_t^{-1}\{f(s)\}]_{t=0}$, the theorem is proved.

5. A series expansion for the solution, $U(x, t)$. The inversion integral solution (3.1) can be expressed in series form by using Theorem A. To show that this is possible, the roots of $G(s) = 0$ must be investigated. In this equation, written in the form

$$\sqrt{s}(C_1 s + C_2 s + K_1 + K_2)\cosh 2\sqrt{s} + [(C_1 s + K_1)(C_2 s + K_2) + s]\sinh 2\sqrt{s} = 0, \tag{5.1}$$

substitute $\sqrt{s} = \alpha + \beta i$. Separating the resulting equation into its real and imaginary parts, we obtain

$$A\cosh 2\alpha\cos 2\beta - B\sinh 2\alpha\sin 2\beta + C\sinh 2\alpha\cos\beta - D\cosh 2\alpha\sin 2\beta = 0,$$

and

$$B\cosh 2\alpha\cos 2\beta + A\sinh 2\alpha\sin 2\beta + D\sinh 2\alpha\cos 2\beta + C\cosh 2\alpha\sin 2\beta = 0,$$

where A, B, C and D involve α, β and the constants of the problem. The determinant of this homogeneous system of equations in $\sin 2\beta$ and $\cos 2\beta$ must vanish because $\sin 2\beta$ and $\cos 2\beta$ cannot vanish simultaneously, so

$$(A^2 + B^2 + C^2 + D^2)\sinh 2\alpha \cosh 2\alpha \\ + (AC + BD)(\cosh^2 2\alpha + \sinh^2 2\alpha) = 0.$$

But $(AC + BD)$ can be reduced to a quantity which is of the same sign as α. Consequently, $\alpha = 0$, and all non-zero roots of equation (5.1) are real and negative. If we substitute $i\beta$ for $\sqrt{s}$, (5.1) becomes

(5.2)

$$(-C_1\beta^2 - C_2\beta^2 + K_1 + K_2)\beta \cos 2\beta + [(-C_1\beta^2 + K_1)(-C_2\beta^2 + K_2) - \beta^2]\sin 2\beta$$

A similar process, applied to (5.2) and its derived equation, can be used to prove that (5.2) has no non-zero double roots. Therefore all non-zero poles of $[u(x,s) - \bar{f}(x)/s]$ are simple, real and negative. It is clear that (5.2) has an infinite number of roots, β_n, and that $\sin 2\beta_n = O(\beta_n^{-1})$; hence these roots lie, for sufficiently large β_n, in the intervals $\frac{1}{2}k\pi - \delta \leqq |\beta_n| \leqq \frac{1}{2}k\pi + \delta$, k being an integer which may differ from n, and δ being any positive number. This shows that the poles of $[u(x,s) - \bar{f}(x)/s]$ are separated by parabolas with vertices at $\frac{1}{2}(k + \frac{1}{2})\pi$:

$$r = \tfrac{1}{4}(k + \tfrac{1}{2})^2\pi^2 \csc^2 \theta/2. \tag{5.3}$$

On these parabolas, $[u(x,s) - \bar{f}(x)/s] = O(s^{-1})$. This can be seen by substituting $s = re^{i\theta}$, where r is given by (5.3), into the formula for $[u(x,s) - \bar{f}(x)/s]$. From this order property and with the results of Section 3, $[u(x,s) - \bar{f}(x)/s]$ is seen to satisfy the conditions of Theorem A. It is easy to find that $A_0 = \bar{f}(x)$ and $A_j = d_j\phi_j(x)$ $(j \neq 0)$, where[6]

$$\phi_j(x) = (K_1 - C_1\beta_j^2)\sin(1 + x)\beta_j + \beta_j \cos(1 + x)\beta_j = -ig_1(1 + x, s_j)$$
$$\psi_j(x) = -ig_2(1 - x, s_j)$$

and

$$d_j = \{\psi_j(-\pi/2\beta_j)[C_1S_1\beta_j^2 + \beta_j \int_{-1}^{1} \phi_j(\xi)f(\xi)\,d\xi] - C_2S_2\beta_j^2\phi(\pi/2\beta_j)\}/\phi_j(\pi/2\beta_j)$$
$$\{[\beta_j^2(C_2\beta_j^2 - K_2)(1 + C_1) + \beta_j^2(C_1\beta_j^2 - K_1)(1 - K_2)]\sin 2\beta_j$$
$$-\tfrac{1}{2}\beta_j[C_2\beta_j^2 + K_2 + C_1\beta_j^2 + K_1 - 2(K_2 - C_2\beta_j^2)(K_1 - C_1\beta_j^2) + 2\beta_j^2]\cos 2\beta_j$$

---

[6] $h_2(1, s_j)\phi_j(x) + h_1(1, s_j)\psi_j(x) = 0.$

Therefore, the solution of the boundary value problem (1. 1)-(1. 6), expressed in series form, is

$$(5.4) \qquad U(x,t) = \sum_{j=0}^{\infty} d_j\phi_j(x)\, e^{-\beta_j^2 t} \qquad (-1 < x < 1, t \geqq 0).$$

6. An expansion theorem for the initial temperature function. It was shown in Section 3 that $[u(x,s) - \bar{f}(x)/s] = O(s^{-3/2})$ where $-1 < x < 1$ and $-\theta_0 \leqq \theta \leqq \theta_0$. Therefore, for any fixed x in that interval, $L_t^{-1}\{u(x,s) - \bar{f}(x)/s\}$ is a continuous function of t and vanishes at $t = 0$. This shows that $U(x,+0) - \bar{f}(x) = U(x,0) - \bar{f}(x) = 0$, or

$$(6.1) \qquad U(x,+0) = \bar{f}(x) \qquad (-1 < x < 1),$$

and condition (1. 4) of our boundary value problem is satisfied. Further, we see that the series (5. 4) must reduce to $\bar{f}(x)$ when $t = 0$; and this fact gives us an expansion theorem for $\bar{f}(x)$.

THEOREM B. *Let $f(x)$ be a piecewise continuous function, and let $f'(x)$ be bounded and integrable. Then, when $-1 < x < 1$, $\bar{f}(x)$ is represented by the series in* (5. 4) *with* $t = 0$.

We may note here that the functions, $i\phi_j(x)$, are characteristic functions of the modified Sturm-Liouville system obtained from the transformed system of Article 2 by placing $f(x) = 0$ and $S_1 = S_2 = 0$.

7. Verification of the solution. The integral form of solution, (3. 1), will be used to prove that the solution satisfies some of the conditions of the problem, while the series form, (5. 4), will be used for the others. The initial temperature condition is already established (see (6. 1)).

In the interval $\frac{1}{2}(a+1) \leqq x \leqq 1$,

$$u(x,s) - f(x)/s - p(x,s) = O(s^{-3/2}) \qquad (-\theta_0 \leqq \theta \leqq \theta_0)$$

uniformly in x, where

$$p(x,s) = \{[S_2 - f(1)][h_1(x,s)g_2(1,s) + g_1(x,s)h_2(1,s)]\}/[sG(s)].$$

Hence ([3], p. 612, Theorem 10) $L_t^{-1}\{u(x,s) - f(x)/s - p(x,s)\}$ exists for any x in this interval and when $t \geqq 0$. Further,

$$(7.1) \qquad \lim_{t\to 0} \lim_{x\to 1} L_t^{-1}\{u(x,s) - f(x)/s - p(x,s)\} = 0.$$

Since $p(x,s) = O(s^{-1}\exp[(2x-2)\sqrt{r}\cos\theta/2]$ when $-\theta_0 \leqq \theta \leqq \theta_0$, a theorem of Carslaw and Jaeger ([1], p. 401, Theorem 2) may be applied to show that $L_t^{-1}\{p(x,s)\}$ is uniformly convergent with respect to x when $x \leqq 1$, and when $t > 0$. Hence $\lim_{x\to 1} L_t^{-1}\{p(x,s)\} = L_t^{-1}\{p(1,s)\}$, and by (7. 1)

$$(7.2) \qquad \lim_{t\to 0} U(1-0,t) - f(1) = S_2 - f(1).$$

The function $U(x,t)$ satisfies, therefore, the boundary condition (1.6). In the same way, by using the interval $-1 \leqq x \leqq \frac{1}{2}(a-1)$, condition (1.5) is shown to be satisfied.

That $U(x,t)$ satisfies the three remaining conditions, (1.1)-(1.3), can be seen by consideration of its series form, (5.4). For sufficiently large values of β_j, $A_j = O(\beta_j^{-1})$. Since each term of the series, (5.4), contains a factor $e^{-\beta_j^2 t}$, the series converges uniformly in x when $t > 0$. Likewise, the series (obtained by termwise differentiation) for $U_x(x,t)$, $U_{xx}(x,t)$ and $U_t(x,t)$ are uniformly convergent when $t > 0$, showing that termwise differentiation is permissible. Rather lengthy, but straightforward, computation will now show that the conditions (1.1)-(1.3) are satisfied by the series form of the solution, thereby completing the verification.

8. **Uniqueness of the solution.** The method used here in proving the uniqueness of the solution is the same as that employed by Churchill ([2], pp. 659-664). Suppose that the boundary value problem, (1.1)-(1.6), has two solutions, $U_1(x,t)$ and $U_2(x,t)$. Then if we set $V(x,t) = U_1(x,t) - U_2(x,t)$, $V(x,t)$ will satisfy conditions (1.1)-(1.6) with $S_1 = S_2 = f(x) = 0$. For certain conditions on $V(x,t)$, the only solution of this system is $V(x,t) = 0$.

THEOREM C. *Let the solution $V(x,t)$ of the homogeneous boundary value problem described above be required to satisfy the following conditions;*

(a) *For each $t > 0$, V, V_x and V_t exist and are continuous in (x,t) when $-1 \leqq x \leqq 1$;*

(b) *For each $t_0 > 0$, $|V|$ is uniformly bounded when $0 \leqq t \leqq t_0$ and $-1 \leqq x \leqq 1$, and $|V_t|$ has the same property for all x in each closed interval containing none of the points $x = -1$, a, 1; also $\int_0^{t_0} |V_x|\,dt$ converges for each x and vanishes as $t_0 \to 0$, uniformly in x when $-1 \leqq x \leqq 1$.*

Then $V(x,t) \equiv 0$ for all $t > 0$.

The proof of Theorem C follows closely that of Churchill's theorem and will not be given here. The task remaining is that of showing that, under certain conditions, the value of $U(x,t)$ we have found satisfies the conditions of Theorem C; and, as a result, that it is unique.

Conditions (a) of Theorem C are satisfied because $U(x,t)$, $U_x(x,t)$ and $U_t(x,t)$ may be expressed as series which are uniformly convergent when $t > 0$ and $-1 \leqq x \leqq 1$. Assume that $f(x)$, $f'(x)$ and $f''(x)$ are piecewise continuous and that $f'''(x)$ is bounded and integrable. For convenience, let $x = a$ be the only point of discontinuity of $f(x)$ and of each of its first two derivatives. This allows us to integrate $I(x,s)$, defined in Article 2, twice by parts. After these integrations by parts, it becomes apparent that $[u(x,s) - f(x)/s$

$-f''(x)/s^2] = O(s^{-5/2})$ uniformly in x, when x is in any closed interval not containing any of the points $x = -1, a, 1$. Since $f(x)$ and $f''(x)$ are bounded, $|U_t(x,t)|$ is uniformly bounded in any such interval.

When $-1 \leqq x \leqq 1$,

$$(8.1) \quad u(x,s) - f(x)/s + \{D_1 \exp[-(1+x)\sqrt{s}] \\ + D_2 \exp[-|x-a|\sqrt{s}] + D_3 \exp[-(1-x)\sqrt{s}]\}/s = O(s^{-3/2}),$$

where D_1, D_2 and D_3 are constants. The inverse transform of $(s^{-1}e^{-y\sqrt{s}})$ is uniformly bounded when $0 \leqq t \leqq t_0$, so we conclude from (8.1) that $|U(x,t)|$ is uniformly bounded when $0 \leqq t \leqq t_0$, $-1 \leqq x \leqq 1$. Similarly, it can be shown that $|U_x(x,t)| \leqq E_1 + E_2/\sqrt{t}$, E_1 and E_2 being constants. Hence, with the stated restrictions on $f(x)$, any solution of the boundary value problem, (1.1)-(1.6), satisfies the conditions of Theorem C and is therefore, unique. Our results are stated as

THEOREM D. *Let $f(x)$ and its first two derivatives be piecewise continuous, and let $f'''(x)$ be bounded and integrable. Then $U(x,t)$, as defined by* (3.1), *is a solution of the boundary value problem,* (1.1)-(1.6); *and it is the only solution of the problem which satisfies the conditions of Theorem C.*

UNIVERSITY OF ALABAMA.

BIBLIOGRAPHY.

1. Carslaw, H. S. and Jaeger, J. C., "A problem in conduction of heat," *Proceedings of the Cambridge Philosophical Society*, vol. 35 (1939), pp. 394-404.

2. Churchill, R. V., "On the problem of temperatures in a non-homogeneous bar with discontinuous initial temperatures," *American Journal of Mathematics*, vol. 61 (1939), pp. 651-664.

3. ———, "The solution of linear boundary value problems in physics by means of the Laplace transformation—Part I. *Mathematische Annalen*, vol. 114 (1937), pp. 591-613.

4. ———, "Additional notes on the inversion of the Laplace transformation," *Mathematische Zeitschrift*, vol. 43 (1938), pp. 743-748.

5. Ince, E. L., *Ordinary Differential Equations*, London (1927).

6. Langer, R. E., "A problem in diffusion or in flow of heat for a solid in contact with a fluid," *Tôhoku Mathematical Journal*, vol. 35 (1932), pp. 260-275.

7. Lowan, A. N., "On conduction of heat in a bar in contact with a stirred liquid," *Philosophical Magazine* (7), vol. 17 (1934), pp. 849-854.

8. March, H. W. and Weaver, Warren, "The diffusion problem for a solid in contact with a stirred liquid," *Physical Review*, vol. 31 (1928), pp. 1072-1082.

9. Peddie, W., "Note on the cooling of a sphere in a mass of well-stirred liquid," *Proceedings of the Edinburgh Mathematical Society*, vol. 19 (1901), p. 34.

10. Peek, R. L., Jr., "Solution to a problem in diffusion employing a non-orthogonal sine series," *Annals of Mathematics* (2), vol. 30 (1929), pp. 265-269.

11. Schumann, T. E. W., "The diffusion problem for a solid in contact with a stirred liquid," *Physical Review* (2), vol. 37 (1931), pp. 1508-1515.

CALCUTTA UNIVERSITY CENTRAL LIBRARY 456

CORRECTION.*

By H. A. RADEMACHER.

Professor D. H. Lehmer has brought to my attention a mistake on pp. 242, 243 of my paper "The Fourier series and the functional equation of the absolute modular invariant $J(\tau)$," this JOURNAL, vol. LXI (1939). It can be corrected as follows: C, the last letter on p. 242, has to be replaced by $O\{k(j^{-1}+(k-j)^{-1})\}$. With this estimate we obtain on p. 243

$$\sum_{m=K+1}^{\infty} \frac{e^{2\pi i j m/k}}{k\tau - m} = O\{kK^{-1}(j^{-1}+(k-j)^{-1})\}.$$

This leads to

$$\begin{aligned} S_1 &= O\{k^{-\frac{1}{3}+\epsilon}K^{-1}\sum_{j=1}^{k-1}(j^{-1}+(k-j)^{-1}\} \\ &= O(k^{-\frac{1}{3}+\epsilon}K^{-1}\log k), \end{aligned}$$

which justifies the estimate given for $T_k^{(K)}$.

UNIVERSITY OF PENNSYLVANIA.

---

* Received March 21, 1942.

VALUATIONS AND INFINITELY NEAR ALGEBRAIC LOCI.*

CALCUTTA CENTRAL LIBRARY UNIVERSITY

By H. T. Muhly.

Introduction. In his paper "Polynomial ideals defined by infinitely near base points," Zariski has shown that the geometric notion of "infinitely near" base points on an algebraic surface can be succinctly expressed in terms of ideals associated with valuations of the field of rational functions determined by the algebraic surface under consideration. By studying the properties of the valuation ideals (v-ideals) in the ring $K[x, y]$ (where K is algebraically closed and of characteristic zero and x and y are indeterminates) which belong to valuations of the field $K(x, y)$, he exposed powerful arithmetic methods for treating questions which arise in connection with base conditions.

In this paper we obtain a generalization of Zariski's results to certain types of higher dimensional base conditions. The central part of the investigation involves proving theorems analogous to those of Zariski, but without the assumption that the ground field K is algebraically closed. Our field K is of characteristic zero, but otherwise arbitrary.

The first four sections of the paper are devoted to a study of the decomposition of the valuations of $K(x, y)$ and of the v-ideals in $K[x, y]$ which occurs when we pass from K to a normal algebraic extension field K^* of K (in particular K^* may be the algebraic closure of K). The fundamental result is given in Theorem 4. 4 and states that Q, the simple ideal of kind h belonging to a valuation B of $K(x, y)$, becomes the product of the distinct conjugate simple ideals of kind h which belong to the extended valuations $\{B^*_\alpha\}$ of B to $K^*(x, y)$. By virtue of this result, most of the theorems valid in $K^*[x, y]$ (when K^* is the algebraic closure of K) carry over to $K[x, y]$. Section **5** is devoted to overcoming the few minor difficulties which arise in this connection.

In Section **6** we study a plane Cremona transformation of the de Jonquières type which behaves with respect to the v-ideals in $K[x, y]$ exactly as would a quadratic transformation if K were algebraically closed. In fact, this de Jonquières transformation is, in a sense, compounded of a set of "conjugate" quadratic transformations.

In the concluding sections we point out the geometric significance of our results. After certain preliminary remarks about the simple subvarieties of an r-dimensional variety V_r, we show that the conditions imposed upon systems of $(r-1)$-dimensional subvarieties of V_r by requiring them to pass through

---

* Received October 10, 1940.

$(r-2)$-dimensional base loci (which are themselves simple subvarieties) either proper or "infinitely near" a proper base variety, can be expressed arithmetically in terms of v-ideals. The de Jonquières transformations of Section 6 are used to construct birational transformations which will separate infinitely near $(r-2)$-dimensional loci into proper varieties.

For the sake of clarity we list below a few definitions and theorems which we shall take for granted in the remainder of the text:

a) By a valuation, B, of a field Σ is meant a homomorphic mapping of Σ (the element 0 excluded) upon an ordered abelian group Γ given by a function $v(a)$, $a \in \Sigma$, $v(a) \in \Gamma$ which obeys the axioms

$$(1) \quad v(a \cdot b) = v(a) + v(b); \qquad (2) \quad v(a+b) \geqq \min(v(a), v(b));$$

$$(3) \quad v(a^*) \neq 0, \text{ for at least one } a^* \text{ in } \Sigma.$$

b) The elements a of Σ for which $v(a) \geqq 0$ form a ring in Σ, called the *valuation ring* of B, which is usually denoted by $\mathfrak{B}$. The elements a of $\mathfrak{B}$ for which $v(a) > 0$ form a prime ideal $\mathfrak{P}$ in $\mathfrak{B}$ which consists of all non-units in $\mathfrak{B}$. The residual ring $\mathfrak{B}/\mathfrak{P}$ is a field and is referred to as the *residue field* of B.

c) If K is a subfield of Σ such that $a(\neq 0) \subset K$ implies $v(a) = 0$, then $\mathfrak{B}/\mathfrak{P}$ contains a field K_1 simply isomorphic with K. The degree of transcendency of $\mathfrak{B}/\mathfrak{P}$ over K_1 is called the dimension of B relative to K.

d) If $\mathfrak{o}$ is a subring of Σ and if $\mathfrak{B} \supset \mathfrak{o}$, then an ideal $\mathfrak{A}$ in $\mathfrak{o}$ of the form $\mathfrak{A} = \mathfrak{M} \cap \mathfrak{o}$, where $\mathfrak{M}$ is an ideal in $\mathfrak{B}$, is called a *v-ideal for B* in $\mathfrak{o}$. A v-ideal $\mathfrak{A}$ for B in $\mathfrak{o}$ is characterized by the property that if $a \equiv 0(\mathfrak{A})$ and b is an element of $\mathfrak{o}$ such that $v(b) \geqq v(a)$ then $b \equiv 0(\mathfrak{A})$. An ideal $\mathfrak{A}$ is referred to simply as a *v-ideal* if it is a v-ideal for some valuation of Σ.

If Hilbert's base condition holds in $\mathfrak{o}$, then, as a varies over the elements of an ideal $\mathfrak{M}$ in $\mathfrak{o}$, $v(a)$ assumes a positive minimum. This minimum is spoken of as the *value of* $\mathfrak{M}$, in symbols $v(\mathfrak{M})$.

e) Under the hypothesis that Hilbert's base condition holds in $\mathfrak{o}$, the v-ideals in $\mathfrak{o}$ for a given valuation B form a well ordered sequence, the first element of which is $\mathfrak{o}$ itself, and the successor of any element $\mathfrak{q}$ consists of those elements a in $\mathfrak{o}$ such that $v(a) > v(\mathfrak{q})$.

For further properties of valuations and of v-ideals we refer to Krull 1 and Zariski 2.

1. Normal ground field extensions. Let Σ be a field, and let K be a subfield of Σ. We assume only that K is of characteristic zero and that it is maximally algebraic in Σ. Throughout our considerations, K will be thought of as the ground field or field of constants. We consider a normal algebraic extension field K^* of K and we construct the field $\Sigma^* = K^*\Sigma$ as described in Zariski 1.[1]

Let B be a valuation of Σ of dimension zero relative to the ground field K, i. e., if $\mathfrak{B}$ is the valuation ring of B in Σ and $\mathfrak{P}$ the ideal of non-units in $\mathfrak{B}$, then $\mathfrak{B} \supset K$ and $\mathfrak{B}/\mathfrak{P}$ is algebraic over K.[2] We observe that the ring $D^* = K^*\mathfrak{B}$ in Σ^*, is integrally dependent on $\mathfrak{B}$. Hence, if B^* is a valuation of Σ^* which coincides with B on Σ (an extension of B to Σ^*) and if $\mathfrak{B}^*$ is the valuation ring of B^*, then $\mathfrak{B}^* \supset D^*$, (Krull 1).

Let us assume for a moment that K^* is a finite normal extension of K. If this is the case, then the relative degrees $[K^* : K]$ and $[\Sigma^* : \Sigma]$ are the same, as are also the Galois groups $\mathfrak{G}(K^* : K)$ and $\mathfrak{G}(\Sigma^* : \Sigma)$, (Zariski 1, Lemma 1). If n is the relative degree $[\Sigma^* : \Sigma]$ then by a well known theorem (Krull 1) there exists at least one, and at most n, distinct extensions, $B^*_1, \cdots B^*_g$, of B to Σ^* $(g \leqq n)$. Moreover, if $\mathfrak{B}^*_i$ is the v-ring of B^*_i in Σ^*, then the intersection $\mathfrak{B}^*_1 \cap \mathfrak{B}^*_2 \cap \cdots \cap \mathfrak{B}^*_g$ is equal to D^*. In fact, we have already seen that $D^* \subset \mathfrak{B}^*_1 \cap \mathfrak{B}^*_2 \cap \cdots \cap \mathfrak{B}^*_g$. To prove the opposite inclusion, let δ be an element of $\mathfrak{B}^*_1 \cap \cdots \cap \mathfrak{B}^*_g$ and let θ be a primitive element of $K^* : K$. As an element in Σ^*, δ can be written in the form

$$\delta = a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1}, \qquad a_i \in \Sigma$$

and the conjugates of δ over Σ are

$$\delta^{(i)} = a_0 + a_1\theta_i + \cdots + a_{n-1}\theta_i^{n-1},$$

where θ_0 $(= \theta), \theta_1, \cdots, \theta_{n-1}$ are the conjugates of θ over K. Since the determinant $| \theta_i^j | \neq 0$, and since each $\delta^{(i)}$ is in $\mathfrak{B}^*_1 \cap \cdots \cap \mathfrak{B}^*_g$, it follows that a_i is also in $\mathfrak{B}^*_1 \cap \cdots \cap \mathfrak{B}^*_g$, $i = 0, \cdots, n-1$. Thus, since a_i is in Σ, we have $a_i \in \mathfrak{B}$ and $\delta \in D^*$.

If $\mathfrak{P}^*_i$ is the ideal of non-units in $\mathfrak{B}^*_i$, and if $\mathfrak{p}^*_i = \mathfrak{P}^*_i \cap D^*$ then it is known (Krull 1) that the quotient ring $D^*_{\mathfrak{p}^*_i}$ coincides with $\mathfrak{B}^*_i$.

We now return to the general case (K^* an arbitrary normal extension of K) and we assert

---

[1] Σ^* is the collection of all finite sums $\xi^* = a^*_1 \alpha_1 + \cdots + a^*_g \alpha_g$, $\alpha_i \subset \Sigma$, $a^*_i \subset K^*$. Two such sums $\xi^* = \sum a^*_i \alpha_i$, $\eta^* = \sum b^*_j \beta_j$ are identified according to the following rule. If $c_1, \cdots, c_r$ is an independent base of the field $K(a^*_i, b^*_j)$ over K, and if $\xi^* = \sum_{i=1}^{r} c_i \lambda_i$, $\eta^* = \sum_{i=1}^{r} c_i \mu_i$, then $\xi^* = \eta^*$ if and only if $\lambda_i = \mu_i$, $i = 1, 2, \cdots, n$.

[2] Strictly speaking, $\mathfrak{B}/\mathfrak{P}$ contains a field isomorphic with K. We identify this field with K.

THEOREM 1.1. *If B^* is any extension of B to Σ^* with valuation ring $\mathfrak{B}^*$ and prime ideal $\mathfrak{P}^*$, and if $\mathfrak{p} = D^* \cap \mathfrak{P}^*$, then the equality $\mathfrak{B}^* = D^*_{\mathfrak{p}^*}$ holds.*

Proof. That $\mathfrak{B}^* \supseteq D^*_{\mathfrak{p}^*}$ is evident. Let ξ^* be an arbitrary element of $\mathfrak{B}^*$,

$$\xi^* = a^*_1\xi_1 + \cdots + a^*_k\xi_k, \qquad a^*_i \subset K^*, \qquad \xi_i \subset \Sigma.$$

Let $\bar{K}$ be the least normal extension of K that contains the field $K(a^*_1, \cdots, a^*_k)$. Let $\bar{\Sigma} = \bar{K}\Sigma$, $\bar{D} = \bar{K}\mathfrak{B}$, $\bar{\mathfrak{B}} = \mathfrak{B}^* \cap \bar{\Sigma}$. Obviously, we have $\xi^* \subset \bar{\mathfrak{B}}$. On the other hand, since $\bar{K}$ is finite over K it follows by the preceding remarks that $\bar{\mathfrak{B}} = \bar{D}_{\bar{\mathfrak{p}}}$, where $\bar{\mathfrak{p}} = \bar{\mathfrak{P}} \cap \bar{D}$, ($\bar{\mathfrak{P}}$ is the ideal of non-units in $\bar{\mathfrak{B}}$). Hence $\xi^* \subset \bar{D}_{\bar{\mathfrak{p}}}$, and since obviously $\bar{D}_{\bar{\mathfrak{p}}} \subset D^*_{\mathfrak{p}^*}$ we conclude that $\mathfrak{B}^* = D^*_{\mathfrak{p}^*}$.

Letting $K_{\mathfrak{P}}$ denote the residue field $\mathfrak{B}/\mathfrak{P}$, and $K^*_{\mathfrak{P}^*}$ the residue field $\mathfrak{B}^*/\mathfrak{P}^*$ we obtain

COROLLARY 1.2. *The field $K^*_{\mathfrak{P}^*}$ is obtained from $K_{\mathfrak{P}}$ by the extension $K \rightarrow K^*$*, i. e.

$$K^*_{\mathfrak{P}^*} = K^* \cdot K_{\mathfrak{P}}.$$

Proof. We know, by Theorem 1.1, that $K^*_{\mathfrak{P}^*} = \mathfrak{B}^*/\mathfrak{P}^* = D^*_{\mathfrak{p}^*}/\mathfrak{p}^* \cdot D^*_{\mathfrak{p}^*} = D^*/\mathfrak{p}^*$. By applying an analogous theorem proved for ideals by Zariski (Zariski 1, Lemma 2) we conclude that $K^*_{\mathfrak{P}^*} = K^* \cdot K_{\mathfrak{P}}$.

THEOREM 1.3. *If $\mathfrak{B}/\mathfrak{P} = K$, then the valuation B has a unique extension to Σ^*.*

Proof. The existence of at least one extension of B to Σ^* can be established by a standard well-ordering process (see e. g. Krull 1). We prove the uniqueness.

The conditions $K_{\mathfrak{P}} = \mathfrak{B}/\mathfrak{P} = K$ and $K \subset \mathfrak{B}$ are sufficient to insure that the extended ideal $\mathfrak{P} \cdot D^*$ is a prime ideal $\mathfrak{p}^*$ (see Zariski 1, Theorem 3). On the other hand, it is obvious that $\mathfrak{p}^* \cap \mathfrak{B} = \mathfrak{P}$ and since D^* depends integrally on $\mathfrak{B}$, this implies that $\mathfrak{p}^*$ is maximal in D^*, (see Krull 2). If $\bar{\mathfrak{p}}$ is any maximal ideal in D^*, then $\bar{\mathfrak{p}} \cap \mathfrak{B}$ is maximal in $\mathfrak{B}$ and therefore $\bar{\mathfrak{p}} \cap \mathfrak{B} = \mathfrak{P}$. It follows that $\mathfrak{p}^*$ is the only maximal ideal in D^*. (By Zariski 1, Theorem 2′, $\mathfrak{P} \cdot D^*$ is the intersection of all those ideals in D^* which lie over $\mathfrak{P}$). This implies, in view of Theorem 1.1, that there is only one extension of B to Σ^*; its valuation ring is $D^*_{\mathfrak{p}^*}$.

Remark. Let Δ be a field of characteristic zero, and let B be any valuation of Δ. It has been proved by Krull (Krull 1) that if Δ^* is a finite normal extension of Δ and if B^*_1 and B^*_2 are any two distinct extensions of B to Δ^*, then there exists an element θ of the Galois group $\mathfrak{G}(\Delta^* : \Delta)$ such that $\theta \times B^*_1$

$= B^*_2$. This result can be extended to the case in which Δ^* is an infinite normal extension of Δ by a simple well-ordering process. In our case then, any two members of the set $\{B^*_\alpha\}$ of all extensions of B to Σ^* are conjugate over Σ.

2. Sequences of valuation ideals. We now restrict ourselves to the case where $\Sigma = K(x, y)$, K is a field of characteristic zero, and x and y are independent indeterminates over K. In this case the condition that K be maximally algebraic in Σ is obviously satisfied.

Let us consider a valuation B of Σ which is zero-dimensional with respect to the field K and which is such that its valuation ring contains the ring $\mathfrak{o} = K[x, y]$. Let $Q_0 \supset Q_1 \supset Q_2 \supset \cdots$ be the sequence of valuation ideals in $\mathfrak{o}$ for the given valuation B. We have $Q_0 = \mathfrak{o}$, $Q_1 = P$, where $P = \mathfrak{o} \cap \mathfrak{P}$. $\mathfrak{P}$ is the ideal of non-units in $\mathfrak{B}$. It is well known that for every integer i, Q_i is a primary zero dimensional ideal with P as its associated prime ideal. Moreover, if the entire sequence of valuation ideals is transfinite, it is completely determined by the simple sequence $Q_0, Q_1, \cdots, Q_n, \cdots$. In fact, only two cases are possible; either the infinite intersection $[Q_0, Q_1, Q_2, \cdots]$ is the zero ideal, or else it is a prime one dimensional ideal P_1. In the first case, all v-ideals for B occur among the zero-dimensional ideals. In the second case, the v-ideals are of the form $P_1{}^m \cdot Q_n$, $m, n = 0, 1, 2, \cdots$, and $P_1{}^{m_1} \cdot Q_{n_1}$ precedes $P_1{}^{m_2} \cdot Q_{n_2}$ if $m_1 < m_2$, or if $m_1 = m_2$ and $n_1 < n_2$. (See Zariski 2).

We consider the residue field $K_{\mathfrak{P}} = \mathfrak{B}/\mathfrak{P}$ of B. This field is an algebraic extension of the field K. We let K^* be any normal extension of K which contains $K_{\mathfrak{P}}$. In this case, the field $\Sigma^* = K^*\Sigma$ is simply the field $K^*(x, y)$. Our interest centers on the class of valuations $\{B^*_\alpha\}$ of Σ^* which induce the given valuation B on Σ. We can associate with each valuation B^*_α a sequence of valuation ideals in the ring $\mathfrak{o}^* = K^*[x, y]$. Let $\mathfrak{q}_{\alpha 0}$ $(= \mathfrak{o}^*)$, $\mathfrak{q}_{\alpha 1}, \mathfrak{q}_{\alpha 2}, \cdots$, $\mathfrak{q}_{\alpha n}, \cdots$ be the sequence of primary zero-dimensional valuation ideals associated with the valuation B^*_α. If α is fixed, and if we consider any ideal $\mathfrak{q}_{\alpha n}$, then there are in $\mathfrak{o}^*$ only a finite number of ideals conjugate to $\mathfrak{q}_{\alpha n}$ with respect to Σ (or K). This is an immediate consequence of the fact that $\mathfrak{q}_{\alpha n}$ has a finite base, each element of which is a polynomial. Obviously, each of the conjugates of $\mathfrak{q}_{\alpha n}$ is a valuation ideal belonging to one of the valuations $\{B^*_\alpha\}$. More precisely, if θ is any element of the Galois group of $\Sigma^* : \Sigma$, then $\theta \times \mathfrak{q}_{\alpha n} = \mathfrak{q}_{\beta n}$ for a suitably chosen β. Conversely, if $\mathfrak{q}_{\alpha n}$ is the n-th ideal in the sequence for B^*_α, $\mathfrak{q}_{\beta n}$ the n-th ideal belonging to B^*_β, and if θ is an element of the Galois group $\mathfrak{G}(\Sigma^* : \Sigma)$ such that $\theta \times B^*_\alpha = B^*_\beta$, then clearly, $\theta \times \mathfrak{q}_{\alpha n} = \mathfrak{q}_{\beta n}$. We thus conclude that even though there may be infinitely many extended valuations B^*_α (and hence infinitely many sequences $\{\mathfrak{q}_{\alpha n}\}$), there are only a finite number of distinct ideals $\mathfrak{q}_{\alpha n}$ for any integer n, and all of them are conjugates.

Let Q be any ideal in the sequence $\{Q_i\}$ in the ring $\mathfrak{o}$, and suppose that for a fixed α, say $\alpha = 1$, the integer n is such that $\mathfrak{o}^*Q \equiv 0(\mathfrak{q}_{1n})$, $\mathfrak{o}^*Q \not\equiv 0(\mathfrak{q}_{1n+1})$. If $\mathfrak{q}_{1n}, \mathfrak{q}_{2n}, \cdots, \mathfrak{q}_{gn}$ are the conjugates of $\mathfrak{q}_{1n}$, then it is clear that $\mathfrak{o}^*Q \equiv 0(\mathfrak{q}_{1n})$ and that $\mathfrak{o}^*Q$ is not contained in any of the ideals which succeed $\mathfrak{q}_{1n}$ in the various sequences to which it belongs.

THEOREM 2.1. *The extended ideal* $\mathfrak{o}^*Q$ *is equal to the intersection* $[\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$.

Proof. Let $\mathfrak{A} = [\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$. We have seen that $\mathfrak{o}^*Q \equiv 0(\mathfrak{A})$. On the other hand, the ideal $\mathfrak{A}$ is invariant under all permutations of the Galois group of $\Sigma^* : \Sigma$, so that $\mathfrak{A}$ is the extended ideal of its contraction with $\mathfrak{o}$, $\mathfrak{A} = \mathfrak{o}^*(\mathfrak{A} \cap \mathfrak{o})$. (Zariski 1, Theorem 2′). If f is any element of $\mathfrak{A} \cap \mathfrak{o}$, then since $f \subset \mathfrak{q}_{1n}$, we have $v_1(f) \geqq v_1(Q)$.[3] (v_1 denotes the value in the valuation B^*_1). Since f is in $\mathfrak{o}$ and since B^*_1 coincides with B on Σ, we conclude that $v(f) \geqq v(Q)$ and hence that $f \equiv 0(Q)$. Thus $(\mathfrak{A} \cap \mathfrak{o}) \equiv 0(Q)$, so that $\mathfrak{o}^*(\mathfrak{A} \cap \mathfrak{o}) \equiv 0(\mathfrak{o}^*Q)$. It follows that $\mathfrak{o}^*Q = [\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$.

THEOREM 2.2. *Let* $\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}$ *be a complete set of conjugate ideals belonging to the sequences* $\{\mathfrak{q}_{\alpha n}\}$. *If* $\mathfrak{A} = [\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$ *there exists an ideal* Q *in the sequence* $\{Q_i\}$ *such that* $\mathfrak{o}^*Q = \mathfrak{A}$.

Proof. The contracted ideal $\mathfrak{q}_{1n} \cap \mathfrak{o}$ is a v-ideal for the valuation B, since the valuation B^*_1 is an extension of B. Thus $\mathfrak{q}_{1n} \cap \mathfrak{o} = Q$, a member of the sequence $\{Q_i\}$. Since the ideals $\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}$ are conjugate over Σ, we have $Q \equiv 0(\mathfrak{q}_{in})$, $i = 1, \cdots, g$. Hence $\mathfrak{A} \cap \mathfrak{o} = Q$. On the other hand, $\mathfrak{A}$ is an invariant ideal so that

$$\mathfrak{A} = \mathfrak{o}^*(\mathfrak{A} \cap \mathfrak{o}).$$

Hence $\mathfrak{A} = \mathfrak{o}^* \cdot Q$.

Remark. In the special case $\mathfrak{B}/\mathfrak{P} = K$, we have seen that if K^* is any normal extension of K, then the valuation B has a unique extension, B^*, to Σ^*. If $\{Q_i\}$ is the sequence of v-ideals for B in $\mathfrak{o}$, and if $\{\mathfrak{q}_i\}$ is the sequence for B^* in $\mathfrak{o}^*$, then as a consequence of Theorems 2.1 and 2.2 we deduce the existence of a 1:1 correspondence between the ideals Q_i and the ideals of $\mathfrak{q}_i$. This correspondence is of such a nature that if Q and $\mathfrak{q}$ are corresponding ideals, then $\mathfrak{q} = \mathfrak{o}^*Q$, $Q = \mathfrak{q} \cap \mathfrak{o}$.[4]

---

[3] Since $\mathfrak{o}^*Q \equiv 0(\mathfrak{q}_{1n})$, $v_1(Q) \geqq v_1(\mathfrak{q}_{1n})$. On the other hand since $\mathfrak{o}^*Q \not\equiv 0(\mathfrak{q}_{1n+1})$, $v_1(Q) < v_1(\mathfrak{q}_{1n+1})$ and hence $v_1(Q) = v_1(\mathfrak{q}_{1n})$.

[4] In the course of the proof of Theorem 2.1 we showed that $Q = \mathfrak{A} \cap \mathfrak{o} = \mathfrak{q} \cap \mathfrak{o}$.

3. A preliminary lemma. In this section we impose two restrictions. We assume that the extended ground field K^* is the *algebraic closure* of K. We use the notation $\bar{K}$ instead of K^* to emphasize this fact. Our second assumption is to the effect that the center P, of the valuation B (of Σ) in the ring $\mathfrak{o}$, is such that $\mathfrak{o}/P = K$. We may therefore assume that $P = \mathfrak{o}(x, y)$. This second assumption implies that all of the valuations $\{\bar{B}_a\}$ (the extensions of B to $\bar{\Sigma}$ $(= \bar{K} \cdot \Sigma)$) are concentric in $\bar{\mathfrak{o}}$ $(= \bar{K}[x, y])$. Their common center is $\bar{\mathfrak{p}} = \bar{\mathfrak{o}}(x, y)$.

Let $\mathfrak{q}$ be a zero dimensional v-ideal in $\bar{\mathfrak{o}}$ belonging to one of the valuations $\bar{B}_a$, let us say $\bar{B}_1$. Let $\mathfrak{q}_1$ $(= \mathfrak{q})$, $\mathfrak{q}_2, \cdots, \mathfrak{q}_g$ be the distinct conjugates of $\mathfrak{q}$ in $\bar{\mathfrak{o}}$. If θ_i is an automorphism of $\bar{\Sigma}$ over Σ such that $\theta_i \times \mathfrak{q} = \mathfrak{q}_i$, $i = 2, \cdots, g$, and if $\bar{B}_i = \theta_i \times \bar{B}_1$, $i = 2, \cdots, g$, then clearly $\mathfrak{q}_i$ is a v-ideal in $\bar{\mathfrak{o}}$ for $\bar{B}_i$. We denote the intersection $[\mathfrak{q}_1, \cdots, \mathfrak{q}_g]$ by $\mathfrak{A}$. The ideal $\mathfrak{A}$, as well as each of the ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_g$, is a primary ideal with $\bar{\mathfrak{p}}$ as associated prime.

It may be assumed that in each of the valuations $\bar{B}_1, \cdots, \bar{B}_g$ the condition $v_i(y) \geqq v_i(x)$ is satisfied. With this in mind, we consider the quadratic transformation $T : x' = x$, $y' = y/x$, which extends the ring $\bar{\mathfrak{o}}$ to the larger ring $\bar{\mathfrak{o}}' = \bar{K}[x', y']$. If ρ is that integer such that $\mathfrak{A} \equiv 0(\bar{\mathfrak{p}}^{\rho})$, $\mathfrak{A} \not\equiv 0(\bar{\mathfrak{p}}^{\rho+1})$ then the extended ideal $\bar{\mathfrak{o}}'\mathfrak{A}$ is of the form $x'^{\rho}\mathfrak{A}'$, $\mathfrak{A}' \not\equiv 0(x')$ (Zariski 2). We let $\mathfrak{q}'_i$ be that ideal of the sequence of v-ideals for $\bar{B}_i$ in $\bar{\mathfrak{o}}'$ which is such that $v_i(\mathfrak{q}'_i) = v_i(\mathfrak{A}')$, $i = 1, \cdots, g$. Since $\mathfrak{A}'$ is invariant under the Galois group of $\bar{K}$ over K, we conclude that the ideals $\mathfrak{q}_1', \cdots, \mathfrak{q}'_g$ are conjugate ideals.[5] We let A' denote the intersection $[\mathfrak{q}'_1, \cdots, \mathfrak{q}'_g]$.

The Lemma of this section is to the effect that $\mathfrak{A}' = A'$.

Proof. Let h be the smallest integer such that $(x')^h A'$ is the extended ideal of an ideal in $\bar{\mathfrak{o}}$; $(x')^h A' = \bar{\mathfrak{o}}'A$, where we may assume that $A = \bar{\mathfrak{o}} \cap (x')^h A'$. The ideal A is a complete ideal, and since it is the inverse transform of A', its system of subforms, $\Omega(A)$, must be of dimension zero.[6]

---

[5] It is conceivable that some of the ideals $\mathfrak{q}'_1, \cdots, \mathfrak{q}'_g$ may coincide. However we assign them distinct subscripts so as to *associate them* with the valuations to which they belong.

[6] If M is any ideal in $\bar{\mathfrak{o}}$, such that $M \equiv 0(\bar{\mathfrak{p}}^{\mu})$, $M \not\equiv 0(\bar{\mathfrak{p}}^{\mu+1})$, then any element f of M is of the form $f = f_{\mu} + f_{\mu+1} + \cdots$, where f_{μ} is a form of degree μ in x and y. Moreover, there exist elements f in M for which the *subform* f_{μ} is not zero. As f varies over the elements of M, its subform f_{μ} generates a linear K-module of finite dimension. It is this module which we call the subform system of M and denote by $\Omega(M)$.

For the properties of complete ideals and their behavior under quadratic transformations the reader is referred to Zariski 2, particularly Sections 4 and 12. These properties will be used in the present section without further reference.

Let $\bar{\mathfrak{q}}_i = T^{-1}(\mathfrak{q}'_i)$, and let λ_i be an integer such that $\bar{\mathfrak{q}}_i \equiv 0(\bar{\mathfrak{p}}^{\lambda_i})$, $\bar{\mathfrak{q}}_i \not\equiv 0(\bar{\mathfrak{p}}^{\lambda_i+1})$. Since the ideals $\bar{\mathfrak{q}}_i'$ are conjugate so also are the ideals $\bar{\mathfrak{q}}_i$, so that we have $\lambda_1 = \lambda_2 = \cdots = \lambda_g$. Let λ denote this common value. It is not difficult to see that $\lambda \leqq h$. If $\lambda > h$, we would have $\bar{\mathfrak{o}}'(\bar{\mathfrak{p}}^{\lambda-h}A) = (x')^{\lambda}A' \subseteq (x')^{\lambda}\mathfrak{q}'_i$. Hence on passing to the contracted ideals we find $\bar{\mathfrak{p}}^{\lambda-h}A \equiv 0(\bar{\mathfrak{q}}_i)$. Now $A \equiv 0(\bar{\mathfrak{p}}^{h})$, $A \not\equiv 0(\bar{\mathfrak{p}}^{h+1})$ so that $\bar{\mathfrak{p}}^{\lambda-h} \cdot A \equiv 0(\bar{\mathfrak{p}}^{\lambda})$, $\bar{\mathfrak{p}}^{\lambda-h} \cdot A \not\equiv 0(\bar{\mathfrak{p}}^{\lambda+1})$. Hence the system of subforms $\Omega(\bar{\mathfrak{p}}^{\lambda-h} \cdot A)$ is contained in the system $\Omega(\bar{\mathfrak{q}}_i)$. This contradicts the fact that $\dim \Omega(\bar{\mathfrak{q}}_i) = 0$. Thus $\lambda \leqq h$.

The inequality $h \leqq \rho$ can be demonstrated in a similar manner. Assume that $h > \rho$. We see that $\bar{\mathfrak{o}}'\bar{\mathfrak{p}}^{h-\rho} \cdot \mathfrak{A} = (x')^{h}\mathfrak{A}' \equiv 0((x')^{h}A')$. Hence $\bar{\mathfrak{p}}^{h-\rho} \cdot \mathfrak{A} \equiv 0((x')^{h}\mathfrak{A}' \cap \bar{\mathfrak{o}})$. In other words we have

$$\bar{\mathfrak{p}}^{h-\rho} \cdot \mathfrak{A} \equiv 0(A), \tag{1}$$

since $(x')^{h}A' \cap \bar{\mathfrak{o}} = A$. Now $\Omega(\bar{\mathfrak{p}}^{h-\rho} \cdot \mathfrak{A})$ and $\Omega(A)$ both consist of forms of degree h. The congruence (1) therefore implies that $\Omega(\bar{\mathfrak{p}}^{h-\rho} \cdot \mathfrak{A}) \subset \Omega(A)$. If $h > \rho$, this contradicts the fact that $\dim \Omega(A) = 0$.

Since $h \leqq \rho$, we have that $(x')^{\rho} \cdot A'$ is the extended ideal of an ideal in $\bar{\mathfrak{o}}$. Now, $(x')^{\rho} \cdot \mathfrak{q}'_i \cap \bar{\mathfrak{o}} = \bar{\mathfrak{p}}^{\rho-\lambda} \cdot \bar{\mathfrak{q}}_i$. Moreover, if $\bar{\mathfrak{q}}_i$ is that v-ideal for $\bar{B}_i$ in $\bar{\mathfrak{o}}$ such that $v_i(\bar{\mathfrak{p}}^{\rho-\lambda} \cdot \bar{\mathfrak{q}}_i) = v_i(\bar{\mathfrak{q}}_i)$, then $\bar{\mathfrak{p}}^{\rho-\lambda}\bar{\mathfrak{q}}_i = [\bar{\mathfrak{q}}_i, \bar{\mathfrak{p}}^{\rho}]$. We thus have $(x')^{\rho} \cdot A' \cap \bar{\mathfrak{o}} = [\bar{\mathfrak{q}}_1, \cdots, \bar{\mathfrak{q}}_g, \bar{\mathfrak{p}}^{\rho}]$. On the other hand, we see that $v_i(\bar{\mathfrak{q}}_i) = v_i(\bar{\mathfrak{p}}^{\rho-\lambda}\bar{\mathfrak{q}}_i) = v_i((x')^{\rho}) + v_i(\mathfrak{q}'_i) = v_i((x')^{\rho}) + v_i(\mathfrak{A}') = v_i(\mathfrak{A})$. Thus $\mathfrak{A} \subseteqq \bar{\mathfrak{q}}_i \subseteqq \mathfrak{q}_i$, and since $\mathfrak{A} \equiv 0(\bar{\mathfrak{p}}^{\rho})$ the equality $[\bar{\mathfrak{q}}_1, \cdots, \bar{\mathfrak{q}}_g, \bar{\mathfrak{p}}^{\rho}] = \mathfrak{A}$ holds. Since $(x')^{\rho} \cdot A'$ is an extended ideal, and since $(x')^{\rho} \cdot A' \cap \bar{\mathfrak{o}} = \mathfrak{A}$, we see that $\bar{\mathfrak{o}}'\mathfrak{A} = (x')^{\rho} \cdot A'$. Hence $(x')^{\rho} \cdot A' = (x')^{\rho}\mathfrak{A}'$, and $A' = \mathfrak{A}'$.

We conclude this section with a remark concerning the second assumption which we made at the outset. We drop the assumption for the moment, and consider an arbitrary zero-dimensional valuation B of Σ with origin P in $\mathfrak{o}$. The extended ideal $\bar{\mathfrak{o}}P$ is equal to the intersection $[\bar{\mathfrak{p}}_1, \cdots, \bar{\mathfrak{p}}_m]$ where $\bar{\mathfrak{p}}_i$ is a prime zero-dimensional ideal in $\bar{\mathfrak{o}}$. If $\mathfrak{o}/P = K_1$, then the number m is the relative degree $[K_1 : K]$, (Zariski 1, Section 17). The various extensions $\{\bar{B}_{\alpha}\}$ of B to $\bar{\Sigma}$ have their origins among the ideals $\bar{\mathfrak{p}}_i$. If we let $\bar{\mathfrak{p}}_i = \bar{\mathfrak{o}}(x - \alpha_1, y - \beta_1)$, then the elements $\alpha_1, \cdots, \alpha_m$, are conjugates, as are also the elements $\beta_1, \cdots, \beta_m$. We may identify the field K_1 with the field $K(\alpha_1, \beta_1)$. The conjugate fields of K_1 over K are then $K_i = K(\alpha_i, \beta_i)$, $i = 1, 2, \cdots, m$. We consider the field $\Sigma_1 = K_1(x, y)$. This field is a subfield of $\bar{\Sigma}$. It is not difficult to see that all of the valuations of the set $\{\bar{B}_{\alpha}\}$ which have origin at $\bar{\mathfrak{p}}_1$ induce the same valuation on the field Σ_1. In fact, if $\bar{B}_{\alpha}$ and $\bar{B}_{\beta}$ are two of the valuations of $\bar{\Sigma}$ with origin at $\bar{\mathfrak{p}}_1$, and if θ is an element of the Galois group $\mathfrak{G}(\bar{\Sigma} : \Sigma)$ such that $\theta \times B_{\alpha} = B_{\beta}$, then $\theta \times \bar{\mathfrak{p}}_1 = \bar{\mathfrak{p}}_1$ and hence $\theta \times \alpha_1 = \alpha_1$,

$\theta \times \beta_1 = \beta_1$. Thus θ leaves each element of K_1 invariant, so that $\bar{B}_\alpha$ and $\bar{B}_\beta$ induce the same valuation on Σ_1. Let B_1 denote their common contracted valuation of Σ_1. It is clear that the set of those members of the class $\{B_\alpha\}$ which have origin at $\bar{\mathfrak{p}}_1$ constitutes the totality of extensions of B_1 to $\bar{\Sigma}$. Similar considerations, applied to the conjugate fields $K_2, \cdots, K_m$, yield the conclusion that the set $\{\bar{B}_\alpha\}$ may be divided into m subsets each of which satisfies the second assumption of this section.

4. The simple ideals and the fundamental lemma. A valuation ideal $\mathfrak{A}$ in a ring $\mathfrak{J}$ is said to be simple if it cannot be written as the product of two ideals both different from the unit ideal; i. e. $\mathfrak{A}$ is simple if and only if $\mathfrak{A} = \mathfrak{B} \cdot \mathfrak{C}$, $\mathfrak{B} \neq \mathfrak{J}$ implies $\mathfrak{C} = \mathfrak{J}$. In the contrary case, $\mathfrak{A}$ is said to be composite. It has been proved by Zariski (Zariski 2, Section 5) that *a composite v-ideal can be written as a product of two v-ideals both different from the unit ideal.* Although this theorem was stated for the case in which the coefficient field is algebraically closed, the proof makes no use of this assumption, and is valid in all cases which come under our consideration.

We again consider the sequence $\{Q_i\}$ of all v-ideals for B in $\mathfrak{o}$. According to Theorem 2. 1, the extended ideal $\bar{\mathfrak{o}}Q_i$ (Q_i is any element of the sequence $\{Q_i\}$) is of the form $[\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$. The ideal $\bar{\mathfrak{o}}Q_i$ is therefore a complete ideal in the sense of Zariski (Zariski 2, Section 12), i. e. $\bar{\mathfrak{o}}Q_i$ can be written as a product of simple v-ideals in $\mathfrak{o}$, say

$$\bar{\mathfrak{o}}Q_i = \mathcal{P}_1^{a_1} \cdot \cdots \cdot \mathcal{P}_k^{a_k} \tag{1}$$

where $\mathcal{P}_j$ is a simple v-ideal in $\bar{\mathfrak{o}}$ belonging to some valuation of $\bar{\Sigma}$. If θ is an automorphism of $\bar{\Sigma}$ over Σ which sends $\mathcal{P}_1$ into one of its conjugates then since $\bar{\mathfrak{o}}Q_i$ is invariant under θ, we have

$$\prod_{i=1}^{k} \mathcal{P}_i^{a_i} = \prod_{i=1}^{k} (\theta \times \mathcal{P}_i)^{a_i}. \tag{2}$$

By the unique factorization theorem for simple v-ideals in $\bar{\mathfrak{o}}$ (Zariski 2, Section 6), we conclude that the product $\mathcal{P}_1^{a_1} \cdot \mathcal{P}_2^{a_2} \cdot \cdots \cdot \mathcal{P}_k^{a_k}$ contains together with each factor $\mathcal{P}_i^{k_i}$ all of its conjugates over Σ. Hence we have

$$\bar{\mathfrak{o}}Q = (\mathcal{P}_{11} \cdot \cdots \cdot \mathcal{P}_{1r_1})^{\lambda_1} (\mathcal{P}_{21} \cdot \cdots \cdot \mathcal{P}_{2r_2})^{\lambda_2} \cdots (\mathcal{P}_{s1} \cdot \cdots \cdot \mathcal{P}_{sr_s})^{\lambda_s}$$

where each parenthesis encloses a complete set of conjugate ideals over Σ. The ideal $\mathcal{P}_{\nu 1} \cdot \cdots \cdot \mathcal{P}_{\nu r_\nu}$ is an invariant ideal (invariant under all automorphisms of $\bar{\Sigma} : \Sigma$) so that $\mathcal{P}_{\nu 1} \cdot \cdots \cdot \mathcal{P}_{\nu r_\nu} = \bar{\mathfrak{o}}\mathfrak{A}_\nu$ where $\mathfrak{A}_\nu$ is an ideal in $\mathfrak{o}$. (Zariski 1, Theorem 2′). Hence, $\bar{\mathfrak{o}}Q = \mathfrak{o}(\prod_{\nu=1}^{s} \mathfrak{A}_\nu^{\lambda_\nu})$. Since the relation

$\bar{\mathfrak{o}}\mathfrak{A} \cap \mathfrak{o} = \mathfrak{A}$ holds for any ideal $\mathfrak{A}$ in $\mathfrak{o}$ (Zariski 1, Theorem 1) we conclude that $Q = \prod_{\nu=1}^{s} \mathfrak{A}_{\nu}^{\lambda_{\nu}}$.

The remarks above establish the following

Lemma 4.1. *The v-ideal Q_i in $\mathfrak{o}$ is simple if and only if the complete ideal $\bar{\mathfrak{o}}Q_i$ involves one and only one complete set of distinct conjugate simple v-ideals in its unique product decomposition, each occurring to the first power only.*

In Zariski 1, a simple v-ideal is said to be of "kind n" if it is the n-th *simple* ideal which occurs in the sequence of v-ideals to which it belongs. If Q_i is a simple v-ideal in $\bar{\mathfrak{o}}$ then the conjugate simple ideals which occur in its product decomposition in $\bar{\mathfrak{o}}$ are all of the same kind, say h. We shall say then that the ideal Q_i in $\mathfrak{o}$ is of kind h. In the light of later theorems it will be seen that this definition of "kind" is the same as that of Zariski.

Theorem 4.2. *If Q is a simple v-ideal for B of kind h in $\mathfrak{o}$, then $\bar{\mathfrak{o}}Q = \mathfrak{P}_1^{(h)} \cdot \cdots \cdot \mathfrak{P}_r^{(h)}$ where $\mathfrak{P}_1^{(h)}, \cdots, \mathfrak{P}_r^{(h)}$ are the distinct conjugate simple v-ideals of kind h which belong to the valuations of the set $\{\bar{B}_a\}$, the extensions of B to $\bar{\Sigma}$.*

Proof. We use an induction with respect to the kind, h, of Q. The theorem is obviously true for kind 1, for then Q is the origin, P, of B, and $\bar{\mathfrak{o}}P = \bar{\mathfrak{p}}_1 \cdot \bar{\mathfrak{p}}_2 \cdot \cdots \cdot \bar{\mathfrak{p}}_m$. As was pointed out at the end of Section 3, the ideals $\bar{\mathfrak{p}}_i$ are the origins of the various valuations $\{\bar{B}_a\}$.

By Theorem 2.1, $\bar{\mathfrak{o}}Q = [\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$ where $\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}$ is a complete set of conjugate ideals. Now in general we shall have $g > m$, so that we contract those ideals among the $\mathfrak{q}_{in}$ which belong to the same prime. We let $\bar{Q}_\nu$ denote the intersection of those ideals among $\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}$ which belong to the prime $\bar{\mathfrak{p}}_\nu$, $\nu = 1, 2, \cdots, m$. We then have

$$\bar{\mathfrak{o}}Q = [\bar{Q}_1, \cdots, \bar{Q}_m] \tag{3}$$

This is the normal primary decomposition of $\bar{\mathfrak{o}}Q$. Since the ideals $\bar{Q}_i$ are two by two relatively prime, the equality

$$[\bar{Q}_1, \cdots, \bar{Q}_m] = \bar{Q}_1 \cdot \bar{Q}_2 \cdot \cdots \cdot \bar{Q}_m \tag{4}$$

holds. Hence it is sufficient to prove that $\bar{Q}_1 = \mathfrak{P}_1^{(h)} \cdot \cdots \cdot \mathfrak{P}_s^{(h)}$, where $\mathfrak{P}_1^{(h)} \cdot \cdots \cdot \mathfrak{P}_s^{(h)}$ are the conjugate simple ideals of kind h belonging to those members of the set $\{\bar{B}_a\}$ which have origin at $\bar{\mathfrak{p}}_1$. (By an obvious Galois argument a similar statement is then true of the ideals $\bar{Q}_2, \cdots, \bar{Q}_m$). Using the notation of Section 3, we pass to the fields K_1 and Σ_1. We are then free to use the second assumption of that section.

We prove our assertion by means of the induction assumption that Theorem 4. 2 is true for simple v-ideals of kind $h-1$. By Theorem 2. 2 (and the remark at the end of Section 3), the ideal $\bar{Q}_1$ is the extension to $\bar{\mathfrak{o}}$ of a v-ideal Q_1 in the ring $\mathfrak{o}_1 = K_1[x, y]$. By Lemma 4. 1, the ideal $\bar{\mathfrak{o}}Q$ involves one and only one complete set of conjugate simple ideals in its decomposition. Hence the ideal $\bar{\mathfrak{o}}Q_1$ involves one and only one complete set of *conjugates with respect to* K_1, in its decomposition. Thus

$$(5) \qquad \bar{\mathfrak{o}}Q_1 = \bar{Q}_1 = \mathfrak{P}^*{}_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}^*{}_s{}^{(h)}$$

where $\mathfrak{P}^*{}_1{}^{(h)}, \cdots, \mathfrak{P}^*{}_s{}^{(h)}$ are simple, conjugate v-ideals in $\bar{\mathfrak{o}}$.

We now apply the quadratic transformation $T_1 \colon x' = x - \alpha_1,\ y' = \dfrac{y-\beta_1}{x-\alpha_1}$ which has its fundamental point at $\bar{\mathfrak{p}}_1$. T_1 extends the rings $\bar{\mathfrak{o}}$ and $\mathfrak{o}_1$ to the rings $\bar{\mathfrak{o}}' = \bar{K}[x', y']$ and $\mathfrak{o}'_1 = K_1[x', y']$. By the Lemma of Section 3, we have

$$(6) \qquad \bar{\mathfrak{o}}'\bar{Q}_1 = x'^{\sigma}[\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f]$$

where $\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f$ are a complete set of conjugate v-ideals in $\bar{\mathfrak{o}}'$ belonging to those valuations $\{\bar{B}_a\}$ with origin at $\bar{\mathfrak{p}}_1$. On the other hand

$$(7) \qquad \bar{\mathfrak{o}}'(\mathfrak{P}^*{}_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}^*{}_s{}^{(h)}) = x'^{\sigma}\mathfrak{P}'{}_1{}^{(h-1)} \cdot \cdots \cdot \mathfrak{P}'{}_s{}^{(h-1)}$$

(Zariski 2, Theorem 11. 1, Corollary 11. 2). We therefore conclude that

$$(8) \qquad [\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f] = \mathfrak{P}'{}_1{}^{(h-1)} \cdot \cdots \cdot \mathfrak{P}'{}_s{}^{(h-1)}.$$

By Theorem 2. 2, the ideal $[\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f]$ is the extended ideal of a v-ideal in $\mathfrak{o}'_1$ belonging to the valuation B_1 of Σ_1, say

$$(9) \qquad [\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f] = \bar{\mathfrak{o}}'Q'_1.$$

Since $\bar{\mathfrak{o}}'Q'_1$ involves only one set of conjugate simple ideals in its decomposition, we conclude that Q'_1 is simple, necessarily of kind $h-1$. By the induction assumption (and the unique factorization theorem in $\bar{\mathfrak{o}}'$) we conclude that $\mathfrak{P}'{}_1{}^{(h-1)}, \cdots, \mathfrak{P}'{}_s{}^{(h-1)}$ belong to valuations of the set $\{\bar{B}_a\}$. This is then true of their inverse transforms $\mathfrak{P}^*{}_1{}^{(h)}, \cdots, \mathfrak{P}^*{}_s{}^{(h)}$.

THEOREM 4. 3. *If* $\mathfrak{A} = \mathfrak{P}_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}_\lambda{}^{(h)}$, *where* $\mathfrak{P}_1{}^{(h)}, \cdots, \mathfrak{P}_\lambda{}^{(h)}$ *are the distinct conjugate simple ideals of kind h belonging to the extensions* $\{\bar{B}_a\}$ *of B to* $\bar{\Sigma}$ *then* $\mathfrak{A} = [\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}]$, *where* $\mathfrak{q}_{in}$ *is that v-ideal for* $\bar{B}_{a_1}$ *(one of the extensions of B to* $\bar{\Sigma}$*) which is such that* $v_{a_1}(\mathfrak{q}_{1n}) = v_{a_1}(\mathfrak{A})$, *and* $\mathfrak{q}_{1n}, \cdots, \mathfrak{q}_{gn}$ *are the conjugates of* $\mathfrak{q}_{1n}$.

Proof. The proof is again carried out by means of an induction with respect to the kind, h, of the ideals $\mathfrak{P}_i^{(h)}$. The theorem is obviously true for simple ideals of kind 1. We assume that the theorem is true for the simple ideals of kind $h' \leqq h - 1$. We remark here that by reasoning entirely analogous to that used in Theorem 4.2 we may assume in addition that all of the ideals $\mathfrak{P}_1^{(h)}, \cdots, \mathfrak{P}_\lambda^{(h)}$ are primary ideals belonging to the same origin $\bar{\mathfrak{p}}_1 = (x - \alpha_1, y - \beta_1)$.

For the first part of our proof we refer to Zariski 2, for the considerations needed here are almost precisely the same as those which Zariski used to prove that a power product of simple v-ideals is an intersection of v-ideals. The only difference is that we make a stronger induction hypothesis and arrive thereby at a stronger conclusion.

On pages 194 and 195 of Zariski 2, an ideal $\mathfrak{A} = \mathfrak{A}_1 \cdot \mathfrak{A}_2 \cdots \cdot \mathfrak{A}_m$ is considered, where $\mathfrak{A}_i$ is a simple v-ideal of kind h_i and all of the ideals $\mathfrak{A}_i$ belong to the same prime $\mathfrak{p}$. Zariski proves that $\mathfrak{A}$ is an intersection of v-ideals by means of an induction with respect to $h = \max(h_1, \cdots, h_m)$. His reduction in kind, from h to $h - 1$, is achieved by means of the quadratic transformation $x' = x$, $y' = y/x$. This procedure leads to the conclusion that the ideal $\mathfrak{A}$ can be written in the form

$$\mathfrak{A} = [\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_k, \mathfrak{p}^\rho]$$

where $\mathfrak{C}_i$ is a v-ideal, and $\rho = \sum_{j=1}^{m} \rho_j$, ($\rho_j$ is that integer such that $\mathfrak{A}_j \equiv 0(\mathfrak{p}^{\rho_j})$, $\mathfrak{A}_j \not\equiv 0(\mathfrak{p}^{\rho_j+1})$.

If we follow this same procedure using the induction assumption that our Theorem 4.3 is true if each of the conjugate factors $\mathfrak{P}_i$ is of kind $(h - 1)$, then we reach the conclusion that

$$\mathfrak{A} = [\mathfrak{q}_{1m}, \cdots, \mathfrak{q}_{gm}, \bar{\mathfrak{p}}^\rho] \tag{10}$$

where $\mathfrak{q}_{1m}, \cdots, \mathfrak{q}_{gm}$ are conjugate v-ideals belonging to the valuations of the set $\{\bar{B}_a\}$, and ρ is that integer such that $\mathfrak{A} \equiv 0(\bar{\mathfrak{p}}_1^\rho)$, $\mathfrak{A} \not\equiv 0(\bar{\mathfrak{p}}_1^{\rho+1})$. From this point, our proof can be easily completed. In fact if $\mathfrak{q}_{km'}$ is that v-ideal for $\bar{B}_{a_k}$ such that $v_{a_k}(\mathfrak{q}_{km'}) = v_{a_k}(\mathfrak{A})$ then $\mathfrak{A} \equiv 0([\mathfrak{q}_{1m'}, \cdots, \mathfrak{q}_{gm'}])$ and since necessarily

$$[\mathfrak{q}_{1m'}, \cdots, \mathfrak{q}_{gm'}] \equiv 0([\mathfrak{q}_{1m}, \cdots, \mathfrak{q}_{gm}])$$

we must have

$$\mathfrak{A} = [\mathfrak{q}_{1m'}, \cdots, \mathfrak{q}_{gm'}, \bar{\mathfrak{p}}_1^\rho]. \tag{11}$$

Our theorem is thus equivalent to the assertion that

$$[\mathfrak{q}_{1m'}, \cdots, \mathfrak{q}_{gm'}] \equiv 0(\bar{\mathfrak{p}}_1^\rho).$$

Suppose that $\zeta = \Sigma a_{\mu\nu}(x - \alpha_1)^{\mu}(y - \beta_1)^{\nu}$ is an element of $\bar{\mathfrak{o}}$ such that $v_{a_k}(\zeta) \geqq v_{a_k}(\mathfrak{A})$, $k = 1, 2, \cdots, g$. We propose to prove that under this condition $\zeta \equiv 0(\bar{\mathfrak{p}}_1{}^{\rho})$. We apply the transformation T_1, $x' = x - \alpha_1$, $y' = (y - \beta_1)/(x - \alpha_1)$, thereby extending the ring $\bar{\mathfrak{o}}$ to the ring $\bar{\mathfrak{o}}' = \bar{K}[x', y']$. $\bar{\mathfrak{o}}' \cdot \mathfrak{A}$ is then of the form $(x')^{\rho}\mathfrak{A}'$, where $\mathfrak{A}' = \mathfrak{P}'_1{}^{(h-1)} \cdot \cdots \cdot \mathfrak{P}'_{\lambda}{}^{(h-1)}$, $\mathfrak{P}'_j{}^{(h-1)} = T_1(\mathfrak{P}_j{}^{(h)})$.[7] By the induction assumption, $\mathfrak{A}' = [\mathfrak{q}'_{1j}, \cdots, \mathfrak{q}'_{gj}]$, where $\mathfrak{q}'_{1j}, \cdots, \mathfrak{q}'_{gj}$ are conjugate v-ideals in $\bar{\mathfrak{o}}'$ belonging to the valuations $\bar{B}_a$, and $v_k(\mathfrak{A}') = v_k(\mathfrak{q}'_{kj})$. In the polynomial ζ we replace $x - \alpha_1$ by x', and $y - \beta_1$ by $x'y'$ and we obtain

$$\zeta(x', y') = (x')^{\sigma}\Sigma a_{\mu\nu}(x')^{\mu+\nu-\sigma}(y')^{\nu} = (x')^{\sigma}\eta(x', y')$$

where $\eta(x', y') \not\equiv 0(x')$. If $\sigma > \rho$, then $\zeta \equiv 0(\bar{\mathfrak{p}}_1{}^{\rho})$ and we are finished. In the case $\sigma \leqq \rho$ we have

$$v_{a_k}(\zeta) = v_{a_k}((x')^{\sigma}) + v_{a_k}(\eta)$$

and $v_{a_k}(\mathfrak{A}) = v_{a_k}((x')^{\rho}) + v_{a_k}(\mathfrak{A}')$ whence we have $v_{a_k}(\eta) \geqq v_{a_k}(\mathfrak{A}')$. Since $\mathfrak{A}' = [\mathfrak{q}'_{1j}, \cdots, \mathfrak{q}'_{gj}]$ we conclude that $\eta \equiv 0(\mathfrak{A}')$. Since $\rho \geqq \sigma$, $(x')^{\rho}\eta$ is an element of $\bar{\mathfrak{o}}$. Moreover, $(x')^{\rho}\eta \equiv 0(\bar{\mathfrak{o}}'\mathfrak{A})$ so that $(x')^{\rho}\eta \equiv 0(\bar{\mathfrak{o}}'\mathfrak{A} \cap \bar{\mathfrak{o}})$, i. e., $(x')^{\rho}\eta \equiv 0(\mathfrak{A})$. As an element of $\mathfrak{A}$ we can write

$$(x')^{\rho}\eta = \phi_{\lambda}(x - \alpha_1, y - \beta_1) + \phi_{\lambda+1}(x - \alpha_1, y - \beta_1) + \cdots.$$

Since the subforms of $\mathfrak{A}$ are of degree $\geqq \rho$ we have $\lambda \geqq \rho$. Dividing by $(x')^{\rho}$ we obtain

$$\eta = \{\phi_{\lambda}(1, y') + x'\phi_{\lambda+1}(1, y') + \cdots\}(x')^{\lambda-\rho}$$

and since $\eta \not\equiv 0(x')$ we conclude that $\lambda = \rho$ and that $\phi_{\lambda}(1, y') \neq 0$. Thus $\eta(0, y') = \phi_{\lambda}(1, y')$. On the other hand

$$\zeta(x', y') = (x')^{\sigma}\{f_{\sigma}(1, y') + x'f_{\sigma+1}(1, y') + \cdots\}$$

so that

$$\eta = f_{\sigma}(1, y') + x'f_{\sigma+1}(1, y') + \cdots$$

Hence $\phi_{\lambda}(1, y') = \eta(0, y') = f_{\sigma}(1, y')$, and this implies $\sigma = \lambda = \rho$.

As an immediate consequence of Theorem 4. 3 we conclude that if $\mathfrak{P}_1{}^{(h)}, \cdots, \mathfrak{P}_{\lambda}{}^{(h)}$ are the distinct simple ideals of kind h which occur among the v-ideals for the valuations $\bar{B}_a$, then their product $\mathfrak{P}_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}_{\lambda}{}^{(h)}$ is the extension to $\bar{\Sigma}$ of a v-ideal for B in $\mathfrak{o}$, say Q. By Lemma 4. 1, Q is necessarily simple. We see, incidentally, that a simple ideal in $\mathfrak{o}$ is of kind h if it is the h-th simple ideal in the valuation sequence to which it belongs. Theorems 4. 2 and 4. 3 can be formulated as one theorem:

---

[7] For the definition of $T_1(\mathfrak{P}_j{}^{(h)})$ we again refer to Zariski 2. See [6].

THEOREM 4.4. *If Q is the simple ideal of kind h for B in $\mathfrak{o}$, then*

$$\bar{\mathfrak{o}}Q = \mathfrak{P}_1^{(h)} \cdot \cdots \cdot \mathfrak{P}_\lambda^{(h)}$$

where $\mathfrak{P}_1^{(h)}, \cdots, \mathfrak{P}_\lambda^{(h)}$ are the distinct simple v-ideals of kind h occuring among the valuations $\bar{B}_a$, and conversely.

We point out that Theorem 4.4 is also true if instead of extending the ground field to $\bar{K}$ we extend K to K^*, where K^* is any normal extension of K which contains the residue field of B. This most general form for Theorem 4.4 (i.e. with $\bar{K}$ replaced by K^*) is proved in the following manner. If $\mathfrak{o}^* = K^*[x, y]$, then since $\bar{\mathfrak{o}}Q = \bar{\mathfrak{o}} \cdot \mathfrak{o}^*Q$, we have, by Zariski 1, Theorem 1, that $\mathfrak{o}^*Q = \bar{\mathfrak{o}} \cdot Q \cap \mathfrak{o}^*$. If $\{B^*_a\}$ is the class of extensions of B to Σ^* $(= K^*(x, y))$ then by Corollary 1.2, the residue field of each of the valuations B^*_a is K^*. Theorem 1.3 therefore implies that there is a 1 : 1 correspondence between the valuations of the class $\{B^*_a\}$ and those of the class $\{\bar{B}_a\}$. The correspondence is such that if $\bar{B}_a$ and B^*_a are a pair of corresponding elements then $\bar{B}_a$ is the unique extension of B^*_a to $\bar{\Sigma}$, and B^*_a is the contraction of $\bar{B}_a$ with Σ^*. Applying Theorem 4.4 to the rings $\mathfrak{o}^*$ and $\bar{\mathfrak{o}}$ we conclude that if $\mathfrak{P}^*_1{}^{(h)}, \cdots, \mathfrak{P}^*_\mu{}^{(h)}$ are the simple ideals of kind h in $\mathfrak{o}^*$ for the valuations $\{B^*_a\}$, then $\mu = \lambda$, and for a suitable choice of indices we have $\bar{\mathfrak{o}}\mathfrak{P}^*_i{}^{(h)} = \mathfrak{P}_i{}^{(h)}$, $\mathfrak{P}_i{}^{(h)} \cap \mathfrak{o}^* = \mathfrak{P}^*_i{}^{(h)}$, $i = 1, 2, \cdots, \lambda$. Hence $\bar{\mathfrak{o}}(\mathfrak{P}^*_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}^*_\lambda{}^{(h)}) = \bar{\mathfrak{o}}Q$ and therefore $\mathfrak{P}^*_1{}^{(h)} \cdot \cdots \cdot \mathfrak{P}^*_\lambda{}^{(h)} = \mathfrak{o}^*Q$ (Zariski 1, Theorem 1).

This theorem is most useful in studying the properties of v-ideals in a ring $K[x, y]$, with an arbitrary coefficient field K.

5. Properties of v-ideals.

As immediate consequences of Theorem 4.4 we obtain the following three theorems.

THEOREM 5.1. (*Unique factorization theorem for simple ideals in $\mathfrak{o}$*): *If $\mathfrak{A}_1^{\alpha_1} \cdot \cdots \cdot \mathfrak{A}_k^{\alpha_k} = \mathfrak{B}_1^{\beta_1} \cdot \cdots \cdot \mathfrak{B}_m^{\beta_m}$, where $\mathfrak{A}_1, \cdots, \mathfrak{A}_k$, $\mathfrak{B}_1, \cdots, \mathfrak{B}_m$ are simple v-ideals in $\mathfrak{o}$, then $k = m$ and for a suitable choice of indices $\mathfrak{A}_i = \mathfrak{B}_i$, $\alpha_i = \beta_i$, $i = 1, 2, \cdots, k$.*

Proof. By Theorem 4.4, $\bar{\mathfrak{o}}\mathfrak{A}_i = \mathfrak{A}_{i1} \cdot \cdots \cdot \mathfrak{A}_{ig_i}$, $\bar{\mathfrak{o}}\mathfrak{B}_i = \mathfrak{B}_{i1} \cdot \cdots \cdot \mathfrak{B}_{ih_i}$, where the $\mathfrak{A}_{ij}$ and $\mathfrak{B}_{ij}$ are simple v-ideals in $\mathfrak{o}$. Since we have

$$\bar{\mathfrak{o}}(\mathfrak{A}_i^{\alpha_1} \cdot \cdots \cdot \mathfrak{A}_k^{\alpha_k}) = \bar{\mathfrak{o}}(\mathfrak{B}_1^{\beta_1} \cdot \cdots \cdot \mathfrak{B}_m^{\beta_m})$$

our theorem follows by the unique factorization theorem in $\bar{\mathfrak{o}}$ (Zariski 2, Theorem 7.1).

THEOREM 5.2. *If $Q^{(a)}$ is a simple v-ideal of kind a for a valuation B of Σ, and if B' is any other valuation of Σ for which $Q^{(a)}$ is a v-ideal, then $Q^{(a)}$ is also the simple v-ideal of kind a for B'.*

Proof. $Q^{(\alpha)}$ is a simple ideal for B', and since its unique decomposition in $\bar{\mathfrak{o}}$ involves simple ideals of kind α, $Q^{(\alpha)}$ is a simple ideal of kind α for $\bar{B}'$.

THEOREM 5.3. *If* $\{Q^{(n)}\}$ *and* $\{Q'^{(n)}\}$ *are the sequences of simple v-ideals in* $\mathfrak{o}$ *belonging to valuations* B *and* B' *of* Σ, *and if for some integer* i *we have* $Q^{(i)} = Q'^{(i)}$, *then also* $Q^{(n)} = Q'^{(n)}$, $n = 1, 2, \cdots, i$.

Proof. Since $Q^{(i)} = Q'^{(i)}$, we have $\bar{\mathfrak{o}}Q^{(i)} = \bar{\mathfrak{o}}Q'^{(i)}$ so that by the unique-factorization theorem in $\bar{\mathfrak{o}}$, the simple ideals of kind i belonging to the extended valuations $\{\bar{B}_a\}$ coincide with those belonging to the valuations $\{\bar{B}'_a\}$, the extensions of B'. By Theorem 6.1 of Zariski 2, it follows that the simple ideals of kind $n < i$ among the two sets $\{\bar{B}_a\}$ and $\{\bar{B}'_a\}$ must also coincide. Our theorem then follows by Theorem 4.4.

Our Theorems 5.1 and 5.3 are the direct analogues of Theorems 7.1 and 6.1 of Zariski 2. The next two theorems are analogous to the Lemma of Zariski 1, Section 12.

THEOREM 5.4. *If* $\mathfrak{A} = Q_1^{a_1} \cdots Q_g^{a_g}$ *is a power product of simple v-ideals in* $\mathfrak{o}$ *belonging to valuations* $B_1, \cdots, B_\nu$ *of* Σ, *then* $\mathfrak{A} = [Q'_1, \cdots, Q'_m]$, *where* $Q'_1, \cdots, Q'_m$ *are v-ideals in* $\mathfrak{o}$.

Proof. By Theorem 4.4 we have $\bar{\mathfrak{o}}\mathfrak{A} = \prod_{j=1}^{g} \prod_{i=1}^{\nu_j} \mathfrak{P}_{ji}^{a_j}$ where $\bar{\mathfrak{o}}\mathfrak{Q}_j = \prod_{i=1}^{\nu_j} \mathfrak{P}_{ji}$. The ideal $\bar{\mathfrak{o}}\mathfrak{A}$ is therefore a complete ideal in the sense of Zariski so that

$$\bar{\mathfrak{o}}\mathfrak{A} = [\mathfrak{q}_1, \cdots, \mathfrak{q}_m]$$

where $\mathfrak{q}_i$ is a v-ideal in $\bar{\mathfrak{o}}$ belonging to a valuation $\bar{B}_i$ of $\bar{\Sigma}$, $i = 1, 2, \cdots, m$. Since $\bar{\mathfrak{o}}\mathfrak{A} \cap \mathfrak{o} = \mathfrak{A}$ (Zariski 1, Theorem 1), we have

$$\mathfrak{A} = [\mathfrak{q}_1 \cap \mathfrak{o}, \cdots, \mathfrak{q}_m \cap \mathfrak{o}].$$

If $Q'_i = \mathfrak{q}_i \cap \mathfrak{o}$ then clearly Q'_i is a v-ideal for the valuation B_i which $\bar{B}_i$ induces on Σ.

THEOREM 5.5. *If* $\mathfrak{A} = [Q'_1, \cdots, Q'_m]$ *where* Q'_i *is a v-ideal in* $\mathfrak{o}$ *then* $\mathfrak{A} = Q_1^{a_1} \cdots Q_s^{a_s}$, *where* $Q_1, \cdots, Q_s$ *are simple v-ideals in* $\mathfrak{o}$.

Proof. We first show that $\bar{\mathfrak{o}}\mathfrak{A}$ is an intersection of v-ideals in $\mathfrak{o}$. Let $B_1, \cdots, B_m$ be valuations of Σ for which the ideals $Q'_1, \cdots, Q'_m$ are v-ideals, and let $\{\bar{B}_a^{(i)}\}$ be the class of extensions of B_i to $\bar{\Sigma}$. Finally, let $\bar{B}^{(i)}$ be a particular member of the class $\{\bar{B}_a^{(i)}\}$. If $\mathfrak{q}_i$ is that v-ideal for $\bar{B}^{(i)}$ such that $\bar{v}_i(\bar{\mathfrak{o}}\mathfrak{A}) = \bar{v}_i(\mathfrak{q}_i)$, ($\bar{v}_i$ denotes value in $\bar{B}^{(i)}$), and if $\mathfrak{q}_{i1} = \mathfrak{q}_i, \mathfrak{q}_{i2}, \cdots, \mathfrak{q}_{ig_i}$ are the conjugates of $\mathfrak{q}_i$, then clearly $\bar{\mathfrak{o}}\mathfrak{A} \equiv 0(A)$ where

$$A = [\mathfrak{q}_{11}, \cdots, \mathfrak{q}_{1g_1}, \cdots, \mathfrak{q}_{m1}, \cdots, \mathfrak{q}_{mg_m}]$$

Moreover, it is immediately evident that $A \cap \mathfrak{o} = \mathfrak{A}$, for an element ω of $\mathfrak{o}$ is an element of $A \cap \mathfrak{o}$ if and only if $v_j(\omega) \geqq v_j(\mathfrak{A})$, $j = 1, 2, \cdots, m$ (v_j denotes value in B_j). Since A is an invariant ideal, we conclude that $\bar{\mathfrak{o}}\mathfrak{A} = A$. Thus $\bar{\mathfrak{o}}\mathfrak{A}$ is a complete ideal so that

$$\bar{\mathfrak{o}}\mathfrak{A} = \mathfrak{P}_1^{a_1} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$$

where $\mathfrak{P}_i$ is a simple ideal in $\bar{\mathfrak{o}}$. Since $\bar{\mathfrak{o}}\mathfrak{A}$ is invariant, it follows as in the proof of Lemma 4.1 that the product $\mathfrak{P}_1^{a_1} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$ is of the form

$$\bar{\mathfrak{o}}\mathfrak{A} = (\mathfrak{P}_{11} \cdot \cdots \cdot \mathfrak{P}_{1r_1})^{\lambda_1} (\mathfrak{P}_{21} \cdot \cdots \cdot \mathfrak{P}_{2r_2})^{\lambda_2} \cdot \cdots \cdot (\mathfrak{P}_{s1} \cdot \cdots \cdot \mathfrak{P}_{sr_s})^{\lambda_s}$$

where each pair of parentheses enclosed a complete set of conjugate simple ideals. By Theorem 4.4, the equality

$$\mathfrak{P}_{i1} \cdot \cdots \cdot \mathfrak{P}_{ir_i} = \bar{\mathfrak{o}}Q_i$$

holds, where Q_i is a simple v-ideal in $\mathfrak{o}$. We thus have that

$$\bar{\mathfrak{o}}\mathfrak{A} = \bar{\mathfrak{o}}(\prod_{i=1}^{s} Q_i^{\lambda_i})$$

so that by Zariski 1, Theorem 1, we conclude that

$$\mathfrak{A} = \prod_{i=1}^{s} Q_i^{\lambda_i}.$$

THEOREM 5.6. *If B and B' are two valuations of $\bar{\Sigma}$ and if $Q^{(n)} = Q'^{(n)}$ where $Q^{(n)}$ ($Q'^{(n)}$) is the n-th simple v-ideal in $\mathfrak{o}$ for B (B'), then the v-ideals in $\mathfrak{o}$ which belong to B and precede $Q^{(n)}$ coincide with those of B' which precede $Q'^{(n)}$.*

Proof. Let $Q \supset Q^{(n)}$ be a v-ideal in $\mathfrak{o}$ belonging to B. We can write

$$Q = Q^{(1)e_1} \cdot Q^{(2)e_2} \cdot \cdots Q^{(n-1)e_{n-1}}$$

where $Q^{(1)}, Q^{(2)}, \cdots, Q^{(n-1)}$ are the simple v-ideals for B and B' which precede $Q^{(n)}$. On extending to $\bar{\mathfrak{o}}$ we have

$$\bar{\mathfrak{o}}Q = (\mathfrak{P}_{11} \cdot \cdots \cdot \mathfrak{P}_{1r_1})^{e_1} \cdot \cdots \cdot (\mathfrak{P}_{n-1,1}, \cdot \cdots \cdot \mathfrak{P}_{n-1,r_{n-1}})^{e_{n-1}}$$

where $\mathfrak{P}_{i1}, \cdots, \mathfrak{P}_{ir_i}$ are the simple v-ideals of kind i which occur among the valuations of the set $\{\bar{B}_a\}$. If $\{\bar{B}'_a\}$ is the set of conjugate extensions of B' to $\bar{\Sigma}$, then since $Q^{(i)} = Q'^{(i)}$, $i = 1, 2, \cdots, n$, the ideals $\mathfrak{P}_{i1}, \cdots, \mathfrak{P}_{ir_i}$ are the simple v-ideals of kind i which occur among the valuations $\bar{B}'_a$. An obvious modification of the proof of Theorem 4.3 leads to the conclusion that $\bar{\mathfrak{o}}Q$ can be written as the intersection of a complete set of conjugate v-ideals belonging to the valuations B'_a. The ideal $\bar{\mathfrak{o}}Q \frown \mathfrak{o}$ is therefore a v-ideal for B', i. e. Q is a v-ideal for B'.

6. Quadratic and de Jonquières Transformations. In this section we investigate certain Cremona transformations of the plane associated with the ring $K[x, y]$. These transformations are of the de Jonquières type, having only a finite number of fundamental points at finite distance and behaving in their vicinity as would a quadratic transformation. We find, in fact, that one of these transformations affects a sequence of v-ideals in $K[x, y]$ (K-arbitrary) exactly as a quadratic transformation affects a sequence of v-ideals in $\bar{K}[x, y]$, where $\bar{K}$ is algebraically closed (Zariski 2, Sections 4 and 5).

Let P be the origin of the valuation B in the ring $\mathfrak{o}$. By means of a non-singular linear transformation in $\mathfrak{o}$, we can choose the independent variables x and y in such a way that if $x \rightarrow \alpha$ and $y \rightarrow \beta$ in the mapping of $\mathfrak{o}$ upon the field $\mathfrak{o}/P$, then both α and β are primitive elements of $\mathfrak{o}/P$ over K. After embedding $\mathfrak{o}/P$ in the algebraic closure $\bar{K}$ of K and forming the ring $\bar{\mathfrak{o}} = \bar{K}[x, y]$, it is clear that the ideal $\bar{\mathfrak{p}} = \bar{\mathfrak{o}}(x - \alpha, y - \beta)$ is a prime ideal in $\bar{\mathfrak{o}}$ lying over P. The conjugates of $\bar{\mathfrak{p}}$ are $\bar{\mathfrak{p}}_i = \bar{\mathfrak{o}}(x - \alpha_i, y - \beta_i)$, where α_1 $(= \alpha)$, $\alpha_2, \cdots, \alpha_n$ and β_1 $(= \beta)$, $\beta_2, \cdots, \beta_n$ are the conjugates of α and β over K. We then have

$$\bar{\mathfrak{o}}P = [\bar{\mathfrak{p}}_1, \cdots, \bar{\mathfrak{p}}_n] = \bar{\mathfrak{p}}_1 \cdot \bar{\mathfrak{p}}_2 \cdot \cdots \cdot \bar{\mathfrak{p}}_n. \tag{1}$$

The integer n is the relative degree of $\mathfrak{o}/P$ over K.

We can assume, without loss of generality, that if $\bar{B}$ is any extension of B to $\bar{\Sigma}$ with origin at $\bar{\mathfrak{p}}_1$, then $\bar{v}(x - \alpha_1) = \bar{v}(y - \beta_1)$. A similar statement is then true of the extended valuations with origin at $\bar{\mathfrak{p}}_i$, $i = 1, 2, \cdots, n$.

We define the two polynomials $f(x)$ and $g(x)$ in $\mathfrak{o}$ as follows

$$\begin{aligned} f(x) &= (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \\ g(x) &= \sum_{i=1}^{n} \frac{\beta_i (x - \alpha_1) \cdots (x - \alpha_{i-1})(x - \alpha_{i+1}) \cdots (x - \alpha_n)}{(\alpha_i - \alpha_1) \cdots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \cdots (\alpha_i - \alpha_n)} . \end{aligned} \tag{2}$$

The coefficients of both $f(x)$ and $g(x)$ are in K. The polynomial $g(x)$ is such that $g(\alpha_i) = \beta_i$, $i = 1, 2, \cdots, n$. Consequently it is clear that $f(x)$ and $y - g(x)$ are both divisible by P. Moreover, a simple calculation will show that $P = \mathfrak{o}(f(x), y - g(x))$. Our assumption that $\bar{v}(x - \alpha_1) = \bar{v}(y - \beta_1)$ implies that $v(f(x)) \leqq v(y - g(x))$, where v denotes value in B. Hence the quotient $(y - g(x))/f(x)$ is in the valuation ring of B.

The de Jonquières transformation, J, in which we are interested, is defined by the equations

$$\begin{aligned} x' &= x & \qquad x &= x' \\ y' &= (y - g(x))/f(x) & \qquad y &= f(x')y' + g(x'). \end{aligned} \tag{3}$$

It is clear from the equations of the transformation that if $\mathfrak{o}' = K[x', y']$, then $\mathfrak{o}' \supset \mathfrak{o} = K[x, y]$. The equations (3) also define a transformation between the planes associated with the rings $\bar{\mathfrak{o}} = \bar{K}[x, y]$ and $\bar{\mathfrak{o}}' = \bar{K}[x', y']$.

Before investigating the relations between the rings $\mathfrak{o}$ and $\mathfrak{o}'$ we study those which exist between $\bar{\mathfrak{o}}$ and $\bar{\mathfrak{o}}'$. For this purpose we introduce the following succession of quadratic transformations. We let T_1 denote the transformation

$$x_1 = x, y_1 = (y - \beta_1)/(x - \alpha_1).$$

T_1 has a fundamental point at the point corresponding to the ideal $\bar{\mathfrak{p}}_1 = \bar{\mathfrak{o}}(x - \alpha_1, y - \beta_1)$. Moreover, if $\bar{\mathfrak{p}}$ is any prime zero-dimensional ideal in $\bar{\mathfrak{o}}$ which does not divide $x_1 - \alpha_1$ and if $\bar{\mathfrak{o}}_1 = \bar{K}[x_1, y_1]$ then $\bar{\mathfrak{o}}_1\bar{\mathfrak{p}}$ is a prime ideal in $\bar{\mathfrak{o}}_1$ and the equality

$$\bar{\mathfrak{o}}\bar{\mathfrak{p}} = \bar{\mathfrak{o}}_{1(\bar{\mathfrak{o}}_1\bar{\mathfrak{p}})}$$

holds (Zariski 3, Section 18). The extended ideal $\bar{\mathfrak{o}}_1\bar{\mathfrak{p}}$ is defined to be the transform of $\bar{\mathfrak{p}}$ by T_1, $T_1(\bar{\mathfrak{p}}) = \bar{\mathfrak{o}}_1\bar{\mathfrak{p}}$. We let $T_1(\bar{\mathfrak{p}}_2) = \bar{\mathfrak{p}}_2^{(1)} = \bar{\mathfrak{o}}_1(x_1 - \alpha_2, y_1 - \beta_2^{(1)})$. We observe that $\bar{v}(x_1 - \alpha_2) \leqq v(y_1 - \beta_2^{(1)})$ where v denotes value in any one of the valuations $\bar{B}_\alpha$. In fact,

$$y - \beta_2 = y_1(x_1 - \alpha_2) + (\alpha_2 - \alpha_1)(y_1 - \beta_2^{(1)}),$$

and since $\bar{v}(y - \beta_2) = \bar{v}(y_1(x_1 - \alpha_2))$, we must have $\bar{v}(y_1 - \beta_2^{(1)}) \geqq \bar{v}(y - \beta_2) = \bar{v}(x_1 - \alpha_2)$. The transformation T_2 is defined by the equations

$$T_2 : x_2 = x_1, y_2 = (y_1 - \beta_2^{(1)})/(x - \alpha_2)$$

If $\bar{\mathfrak{o}}_2 = K[x_2, y_2]$, then we have $\bar{\mathfrak{o}} \subset \bar{\mathfrak{o}}_1 \subset \bar{\mathfrak{o}}_2$.

Suppose that we have defined step by step the transformations $T_1, T_2, \cdots T_i$. Associated with T_j is the ring $\bar{\mathfrak{o}}_j = \bar{K}[x_j, y_j]$, $j = 1, 2, \cdots, i$, and the relations

$$\bar{\mathfrak{o}} \subset \bar{\mathfrak{o}}_1 \subset \cdots \subset \bar{\mathfrak{o}}_i$$

hold. We operate on $\bar{\mathfrak{p}}_{i+1}$ by each of the T's in turn, and we obtain

$$T_1(\bar{\mathfrak{p}}_{i+1}) = \bar{\mathfrak{p}}_{i+1}^{(1)}, \cdots, T_i(\bar{\mathfrak{p}}_{i+1}^{(i-1)}) = \mathfrak{p}_{i+1}^{(i)},$$

$\bar{\mathfrak{p}}_{i+1}^{(i)} = \bar{\mathfrak{o}}_i(x_i - \alpha_{i+1}, y_i - \beta_{i+1}^{(i)})$. Again, by Zariski 3, Section 18, we have

$$\bar{\mathfrak{o}}\bar{\mathfrak{p}}_{i+1} = \bar{\mathfrak{o}}_i\bar{\mathfrak{p}}_{i+1}^{(i)}. \tag{4}$$

By a straightforward calculation we find that

$$y - \beta_{i+1} = \phi_i(x_i)y_i + \psi_i(x_i); \qquad x - \alpha_{i+1} = x_i - \alpha_{i+1} \tag{5}$$

where $\phi_i(x_i)$ and $\psi_i(x_i)$ are polynomials in x_i, and $\phi_i(\alpha_{i+1}) \neq 0$. We must then have $\beta^{(i)}_{i+1} = -\psi_i(\alpha_{i+1})/\phi_i(\alpha_{i+1})$. Equations (5) therefore show that in any of the valuations $\bar{B}_a$, $\bar{v}(x_i - \alpha_{i+1}) \leqq \bar{v}(y_i - \beta^{(i)}_{i+1})$. We then define T_{i+1} by the equations

$$x_{i+1} = x_i,\ y_{i+1} = (y_i - \beta^{(i)}_{i+1})/(x_i - \alpha_{i+1})$$

and we let $\bar{\mathfrak{o}}_{i+1}$ denote the ring $\bar{K}[x_{i+1}, y_{i+1}]$. We have thus defined the succession of transformations $T_1, T_2, \cdots, T_n$.

LEMMA 6.1. *The ring $\bar{\mathfrak{o}}_n$ coincides with the ring $\bar{\mathfrak{o}}'$.*

Proof. From the definition of y_n given above it follows immediately that

$$y = y_n f(x) + h(x) \tag{6}$$

where $h(x)$ is a polynomial in x of degree $\leqq n - 1$. Since $y = \beta_i$ when $x = \alpha_i$, and since $f(\alpha_i) = 0$, we see from (6) that $h(\alpha_i) = \beta_i$, $i = 1, 2, \cdots, n$. We conclude that $h(x) \equiv g(x)$ since $g(x)$ is the unique polynomial of degree $\leqq n - 1$ which takes the value β_i when $x = \alpha_i$. Hence $y = y_n f(x) + g(x)$, so that by (3), $y_n = y'$.

We have already observed that the origin P of B in $\mathfrak{o}$ is of the form $P = \mathfrak{o}(f(x), y - g(x))$. If A is any ideal in $\mathfrak{o}$, and h is an integer such that $A \equiv 0(P^h)$, $A \not\equiv 0(P^{h+1})$, then any element a in A is of the form

$$a = \alpha_0 f^h + \alpha_1 f^{h-1}(y - g) + \cdots + \alpha_h (y - g)^h. \tag{7}$$

Since $y - g(x) = f(x')y'$ it follows that a is divisible by $(f(x'))^h$ in $\mathfrak{o}'$. Hence $\mathfrak{o}'A = f^h A'$ where A' is an ideal in $\mathfrak{o}'$. Moreover, if $a \not\equiv 0(P^{h+1})$ then it is not difficult to see that a is not divisible by f^{h+1} in $\mathfrak{o}'$. In fact, since $a \not\equiv 0(P^{h+1})$, not all of the coefficients $\alpha_0, \alpha_1, \cdots, \alpha_n$ are in P. If we let $\bar{\alpha}_i(x) = \alpha_i(x, g(x))$, and if we put

$$\bar{a} = \bar{\alpha}_0 f^h + \bar{\alpha}_1 f^{h-1}(y - g) + \cdots + \bar{\alpha}_n (y - g)^h,$$

then $a \equiv \bar{a}(P^{h+1})$ and $a - \bar{a}$ is divisible by f^{h+1} in $\mathfrak{o}'$. In $\mathfrak{o}'$ we have

$$\bar{a} = (\bar{\alpha}_0 + \bar{\alpha}_1 y' + \cdots + \bar{\alpha}_n y'^h) f^h.$$

Since $\bar{\alpha}_i$ is a polynomial in x', it follows that if $\bar{\alpha}_0 + \bar{\alpha}_1 y' + \cdots + \bar{\alpha}_n y'^h$ is divisible by $f(x')$, then $\bar{\alpha}_i \equiv 0(f)$, $i = 0, 1, 2, \cdots, h$. This would imply that $\alpha_i \equiv 0(P)$, $i = 0, 1, 2, \cdots, h$, and this is not the case. Hence a is not divisible by f^{h+1} in $\mathfrak{o}'$ and consequently neither is $\bar{a}$. We conclude that $\mathfrak{o}'A = f^h A'$, $A' \not\equiv 0(f)$. The ideal A' is defined to be the transform of A by the de Jonquières transformation J; in symbols, $A' = J(A)$.

THEOREM 6.2. *If Q is a v-ideal in $\mathfrak{o}$ for the valuation B, then its transform $J(Q)$ is a v-ideal in $\mathfrak{o}'$ for B.*

Proof. We pass to the rings $\bar{\mathfrak{o}}$ and $\bar{\mathfrak{o}}'$. It is easy to see that $\bar{\mathfrak{o}}'Q \cap \mathfrak{o}' = \mathfrak{o}'Q$. In fact, it is obvious that $\bar{\mathfrak{o}}'Q \cap \mathfrak{o}' \supseteqq \mathfrak{o}'Q$. On the other hand, if a is any element of $\bar{\mathfrak{o}}'Q \cap \mathfrak{o}'$ then we have

$$a = b_1 q_1 + \cdots + b_r q_r$$

where $q_1, \cdots, q_r$ are in Q and $b_1, \cdots, b_r$ are elements of $\bar{\mathfrak{o}}'$. If $c_1, \cdots, c_s$ are those elements of K which occur as coefficients among the polynomials $b_1, \cdots, b_r$, and if θ is a primitive element of $K(c_1, \cdots, c_s)$ over K, then we can write

$$b_i = \bar{b}_i(x', y'; \theta), \quad i = 1, 2, \cdots, r.$$

If $\theta_1, \cdots, \theta_m$ $(\theta_1 = \theta)$ are the conjugates of θ we have

$$(8) \qquad a = \bar{b}_1(x', y'; \theta_i) q_1 + \cdots + \bar{b}_r(x', y'; \theta_i) q_r .$$

On adding the equations (8) we have

$$ma = B_1(x', y') q_1 + \cdots + B_r(x', y') q_r$$

where $B_j = \sum_{i=1}^{m} \bar{b}_j(x', y'; \theta_i)$. Obviously the polynomial B_j is in $\mathfrak{o}'$, and consequently we have $a \equiv 0(\mathfrak{o}'Q)$. This proves that $\mathfrak{o}'Q \cap \bar{\mathfrak{o}}' = \mathfrak{o}'Q$.

To determine the behavior of Q as it is extended to $\bar{\mathfrak{o}}'$, we note first that by Theorem 2.1, $\bar{\mathfrak{o}}Q = [\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_g]$ where $\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_g$ comprise a complete set of conjugate v-ideals belonging to the extensions $\{\bar{B}_\alpha\}$ of B to $\bar{\Sigma}$. We can write $[\mathfrak{q}_1, \cdots, \mathfrak{q}_g] = [\bar{Q}_1, \bar{Q}_2, \cdots, \bar{Q}_n]$ where $\bar{Q}_i$ is a primary ideal with $\bar{\mathfrak{p}}_i$ as its associated prime ideal. $\bar{Q}_i$ is the intersection of those ideals among $\mathfrak{q}_1, \cdots, \mathfrak{q}_g$ which belong to valuations with center at $\bar{\mathfrak{p}}_i$. We thus have

$$(9) \qquad \bar{\mathfrak{o}}Q = \bar{Q}_1 \cdot \bar{Q}_2 \cdots \bar{Q}_n$$

We can obtain the extended ideal $\bar{\mathfrak{o}}'Q$ by extending $\bar{\mathfrak{o}}Q$ step by step through the sequence of rings $\bar{\mathfrak{o}}_1, \bar{\mathfrak{o}}_2, \cdots, \bar{\mathfrak{o}}_N$. If h is that integer such that $Q \equiv 0(P^h)$, $Q \not\equiv 0(P^{h+1})$ then obviously, $\bar{Q}_i \equiv 0(\bar{\mathfrak{p}}_i^h)$, $\bar{Q}_i \not\equiv 0(\bar{\mathfrak{p}}_i^{h+1})$. By the Lemma of Section **3**, we have that $\bar{\mathfrak{o}}_1\bar{Q}_1 = (x_1 - \alpha_1)^h Q_1^{(1)}$, where $Q_1^{(1)}$ is an intersection of conjugate v-ideals belonging to those valuations in the set $\{\bar{B}_\alpha\}$ which have origin at $\bar{\mathfrak{p}}_1$. On the other hand, $\bar{\mathfrak{o}}_1\bar{Q}_i = \bar{Q}_i^{(1)}$, $i = 2, \cdots, n$. Since $\bar{\mathfrak{p}}_i$ is not a fundamental point of T_1 if $i \geqq 2$, the ideal $Q_i^{(1)}$ is also an intersection of conjugate v-ideals in $\bar{\mathfrak{o}}_1$ and $Q_i^{(1)}$ induces the same ideal in the quotient ring $\bar{\mathfrak{o}}_{1\mathfrak{p}_i^{(1)}} = \bar{\mathfrak{o}}_{\mathfrak{p}_i}$ as does Q_i.

As we continue this process, it is clear that when the prime ideal associated with $Q_i^{(j)}$ is not a fundamental point of T_k then $Q_i^{(j)}$ is essentially unaltered as it is extended from $\bar{\mathfrak{o}}_k$ to $\bar{\mathfrak{o}}_{k+1}$. Again, if $\bar{Q}_i^{(j-1)}$ is associated with the prime

$\bar{\mathfrak{p}}_i^{(i-1)}$ in the ring $\bar{\mathfrak{o}}_{i-1}$, then $\bar{\mathfrak{o}}_i \cdot \bar{Q}_i^{(i-1)} = (x_i - \alpha_i)^h \bar{Q}_i^{(i)}$. Thus, in n steps we conclude that

$$\bar{\mathfrak{o}}_n Q = \Big(\prod_{i=1}^{n} (x - \alpha_i)^h\Big) A = f^h A,$$

where A is an intersection of sets of conjugate v-ideals, each set consisting of ideals belonging to valuations with a common center in $\bar{\mathfrak{o}}$. Since A is obviously an invariant ideal these various sets must also be conjugate, one to another. By Theorem 2.2, A is therefore the extension to $\bar{\mathfrak{o}}'$ of a v-ideal Q' in $\mathfrak{o}'$ belonging to the valuation B, and $A \cap \mathfrak{o}' = Q'$. Hence we have $f^h A \cap \mathfrak{o}' = f^h Q'$ and consequently $\mathfrak{o}' Q = f^h Q'$, $J(Q) = Q'$.

THEOREM 6.3. *If $Q^{(k)}$ is the simple ideal of kind k in $\mathfrak{o}$ for the valuation B, then $J(Q^{(k)}) = Q'^{(k-1)}$, the simple ideal of kind $k-1$ for B in $\mathfrak{o}'$.*

Proof. Theorem 6.3 is proved in much the same way as was Theorem 6.2. By Theorem 4.4,

$$\bar{\mathfrak{o}} Q^{(k)} = \mathfrak{P}_1^{(k)} \cdot \cdots \cdot \mathfrak{P}_r^{(k)}$$

where $\mathfrak{P}_1^{(k)}, \cdots, \mathfrak{P}_r^{(k)}$ are the distinct conjugate simple ideals of kind k which occur among the valuations $\{\bar{B}_\alpha\}$. As we extend $\bar{\mathfrak{o}} Q^{(k)}$ step by step through the sequence $\bar{\mathfrak{o}}_1, \bar{\mathfrak{o}}_2, \cdots, \bar{\mathfrak{o}}_n$ each ideal $\mathfrak{P}_i^{(k)}$, $i = 1, \cdots, r$, either remains of kind k or is transformed into an ideal of kind $k-1$. In n steps we find that

$$\bar{\mathfrak{o}}_n Q^{(k)} = f^h (\mathfrak{P}_i'^{(k-1)} \cdot \cdots \cdot \mathfrak{P}_r'^{(k-1)}),$$

where $\mathfrak{P}_i'^{(k-1)}, \cdots, \mathfrak{P}_r'^{(k-1)}$ are the conjugate simple ideals in $\bar{\mathfrak{o}}_n$ $(= \bar{\mathfrak{o}}')$ which occur among the valuations $\{\bar{B}_\alpha\}$. By Theorem 4.4, we have $\bar{\mathfrak{o}}_n Q^{(k)} \cap \mathfrak{o}' = f^h Q'^{(k-1)}$, where $Q'^{(k-1)}$ is the simple ideal of kind $k-1$ in $\mathfrak{o}'$ belonging to B. Hence $J(Q^{(k)}) = Q'^{(k-1)}$.

THEOREM 6.4. *If Q' is a v-ideal in $\mathfrak{o}'$ belonging to the valuation B, then there exists an integer h and a v-ideal Q in $\mathfrak{o}$ such that the relations $\mathfrak{o}' Q = f^h Q'$, and $f^h Q' \cap \mathfrak{o} = Q$ hold.*

Proof. The proof is more or less a reversal of that of Theorem 6.2. We can write

$$\bar{\mathfrak{o}}' Q' = [\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f] = \bar{Q}'_1 \cdot \cdots \cdot \bar{Q}'_n$$

where $\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f$ are conjugate v-ideals in $\bar{\mathfrak{o}}'$ belonging to the valuations $\{\bar{B}_\alpha\}$, and $\bar{Q}'_i$ is the intersection of those ideals among $\mathfrak{q}'_1, \cdots, \mathfrak{q}'_f$ which belong to valuations which have center at $\bar{\mathfrak{p}}_i$ in $\bar{\mathfrak{o}}$. Evidently, the relation $\bar{\mathfrak{o}}' q_1 = \bar{Q}'_1$ holds, where $q_1 = \bar{Q}'_1 \cap \bar{\mathfrak{o}}_1$. By Zariski 2, Theorem 4.3, there is an integer h such that $(x_1 - \alpha_1)^h q_1$ is the extended ideal of an ideal in $\bar{\mathfrak{o}}$. Assuming that

h is the smallest integer with this property, we let $\bar{Q}_1$ denote the ideal $\bar{\mathfrak{o}} \cap (x_1 - \alpha_1)^h q_1$. We then have $\bar{\mathfrak{o}}'\bar{Q}_1 = (x' - \alpha_1)^h \bar{Q}'_1$. If $\bar{Q}_2, \cdots, \bar{Q}_n$ are the conjugates of $\bar{Q}_1$ it is immediately evident that

$$\bar{\mathfrak{o}}'(\bar{Q}_1 \cdots \cdots \bar{Q}_n) = f^h(\bar{Q}'_1 \cdots \cdots \bar{Q}'_n), \quad f^h(\bar{Q}'_1 \cdots \cdots \bar{Q}'_n) \cap \bar{\mathfrak{o}} = \bar{Q}_1 \cdots \cdots \bar{Q}_n.$$

By Theorem 2. 2, the ideal $\bar{Q}_1 \cdots \cdots \bar{Q}_n$ is the extension to $\bar{\mathfrak{o}}$ of a v-ideal Q in $\mathfrak{o}$. It follows at once that $\mathfrak{o}'Q = f^h Q'$, $f^h Q' \cap \mathfrak{o} = Q$.

Remark. One can verify in a similar manner that if k is an integer and $k \geqq h$, then $f^k Q' \cap \mathfrak{o} = P^{k-h} \cdot Q$. Moreover, $\mathfrak{o}'\,(P^{k-h} \cdot Q) = f^k Q'$. However, the ideal $P^{k-h}Q$ need not be a v-ideal.

Theorems 6. 2, 6. 3, 6. 4 show that the behavior of the v-ideals in $\mathfrak{o}$ under the transformation J is in all respects exactly like the behavior of the v-ideals in a ring with an algebraically closed coefficient field under a quadratic transformation (Zariski 2).

7. The simple subvarieties of an algebraic variety. Let V_r be an irreducible r-dimensional algebraic variety in the affine space S_n. We assume that the underlying field of constants, K, is algebraically closed and of characteristic zero. If $(\xi_1, \cdots, \xi_n)$ is the general point of V_r in S_n, then $\Sigma = K(\xi_1, \cdots, \xi_n)$ is the field of rational functions on V_r. With the particular model V_r of Σ we have associated the ring $\mathfrak{J} = K[\xi_1, \cdots, \xi_n]$. It is permissible to assume that $\xi_1, \cdots, \xi_r$ are algebraically independent and that $\xi_{r+1}, \cdots, \xi_n$ depend integrally on $\xi_1, \cdots, \xi_r$. This we do.

With each subvariety of V_r (not on the hyperplane at infinity) there is associated an ideal in $\mathfrak{J}$ and with each ideal there is associated a subvariety of V_r. In particular, if V_s is an irreducible subvariety of V_r it is associated with a prime ideal in $\mathfrak{J}$ of dimension s, and conversely. Zariski calls a variety $V_s(\mathfrak{p}_s)$ a simple subvariety of V_r if there exist $r - s$ algebraically independent elements $\eta_1, \cdots, \eta_{r-s}$ in $\mathfrak{J}$ such that

$$\mathfrak{J}_{\mathfrak{p}_s}(\eta_1, \cdots, \eta_{r-s}) = \mathfrak{J}_{\mathfrak{p}_s}\mathfrak{p}_s,$$

where $\mathfrak{J}_{\mathfrak{p}_s}$ denotes the quotient ring of $\mathfrak{J}$ at $\mathfrak{p}_s$ (Zariski 1). He also proves that in order that $V_s(\mathfrak{p}_s)$ be a simple subvariety of V_r it is necessary and sufficient that there exist an element ω in $\mathfrak{J}$ such that $\dfrac{\partial f}{\partial \omega} \not\equiv 0(\mathfrak{p}_s)$, where $f(\xi, \omega) = \omega^\nu + \alpha_1(\xi_1, \cdots, \xi_r)\omega^{\nu-1} + \cdots + \alpha_\nu(\xi_1, \cdots, \xi_r) = 0$ is the equation expressing the integral dependence of ω on $\xi_1, \cdots, \xi_r$. It follows readily from this last condition that on every simple subvariety of V_r there is at least one simple point of V_r. Elements such as $\eta_1, \cdots, \eta_{r-s}$ defined above are called uniformizing parameters along V_s.

LEMMA 7.1. *If $\mathfrak{J}$ defines a simple subvariety of V_r, and if $\mathfrak{p}_0 \supset \mathfrak{p}_s$ is a zero-dimensional prime ideal in $\mathfrak{J}$ which defines a simple point of V_r on V_s, then it is possible to find elements $\eta_1, \cdots, \eta_r$ in $\mathfrak{J}$ satisfying the conditions*

a) $\eta_1, \cdots, \eta_r$ *are unformizing parameters at* $\mathfrak{p}_0$.

b) $\mathfrak{p}_0 = (\eta_1, \cdots, \eta_r; \mathfrak{p}_s)$.

c) *Every element of $\mathfrak{J}$ is integrally dependent on* $K[\eta_1, \cdots, \eta_r]$.

Proof. We may assume that $\xi_1, \cdots, \xi_r$ are uniformizing parameters at $\mathfrak{p}_0$ and that $\mathfrak{p}_0 = \mathfrak{J}(\xi_1, \cdots, \xi_n)$, (Zariski 4). We let $\eta^*_i = u_{i1}\xi_1 + \cdots + u_{in}\xi_n$, $i = 1, 2, \cdots, r$ where the u's are indeterminates. Since $s < r$, it is clear that for generic values of the u_{ij} the ideal $\mathfrak{J}(\eta^*_1, \cdots, \eta^*_{r-1}, \mathfrak{p}_s)$ will be zero dimensional. Hence for generic values of the u_{ij} the ideal $\mathfrak{J}(\eta^*_1, \cdots, \eta^*_r, \mathfrak{p}_s)$ will not be divisible by a prime ideal in $\mathfrak{J}$ other than $\mathfrak{p}_0$. On the other hand it is known that for non-special values of the u_{ij}, conditions a) and c) will be satisfied by $\eta^*_1, \cdots, \eta^*_r$. We can therefore choose constant values for the u_{ij} say $\bar{u}_{ij}$ so that the quantities $\eta_1, \cdots, \eta_r$ obtained from $\eta^*_1, \cdots, \eta^*_r$ by putting $u_{ij} = \bar{u}_{ij}$ will satisfy conditions a, b, and c.

We consider any ideal $\mathfrak{p}_s$ which defines a simple subvariety of V_r, and we fix a simple point $\mathfrak{p}_0$ on $\mathfrak{p}_s$, $\mathfrak{p}_0 \supset \mathfrak{p}_s$. We assume that the independent variables $\xi_1, \cdots, \xi_r$ satisfy conditions a, b and c of Lemma 7.1, and we denote the ring $K[\xi_1, \cdots, \xi_r]$ by R. Since $\mathfrak{J}$ depends integrally on R, the ideal $P_s = \mathfrak{p}_s \cap R$ is a prime s-dimensional ideal in R. We let $\mathfrak{M}$ and $\mathfrak{N}$ denote the quotient rings R_P, and $\mathfrak{J}_\mathfrak{p}$, respectively, and we put $P^* = \mathfrak{M} \cdot P_s$, $\mathfrak{p}^* = \mathfrak{N} \cdot \mathfrak{p}_s$.

THEOREM 7.2. *There is a* 1:1 *correspondence between the primary ideals Q^* in $\mathfrak{M}$ which belong to P^* and the primary ideals $\mathfrak{q}^*$ in $\mathfrak{N}$ which belong to $\mathfrak{p}^*$. The correspondence is of such a nature that if Q^* and $\mathfrak{q}^*$ are a corresponding pair, then $Q^* = \mathfrak{q}^* \cap \mathfrak{M}$, $\mathfrak{q}^* = \mathfrak{N} \cdot Q^*$.*

Proof. Let $\mathfrak{p}^*$ be a primary ideal in $\mathfrak{N}$ belonging to $\mathfrak{p}^*$, and let $\mathfrak{q}^* \cap \mathfrak{J} = \mathfrak{q}$. The ideal $\mathfrak{q}$ is primary, and $\mathfrak{p}_s$ is its associated prime. It is clear that

$$\mathfrak{p}_0 = (\xi_1, \cdots, \xi_r, \mathfrak{q}), \tag{1}$$

for any prime ideal which divides $(\xi_1, \cdots, \xi_r, \mathfrak{q})$ must divide $(\xi_1, \cdots, \xi_r, \mathfrak{p}_s) = \mathfrak{p}_0$. Moreover, no primary ideal belonging to $\mathfrak{p}_0$ divides $(\xi_1, \cdots, \xi_r)$. If $(\xi_1, \cdots, \xi_r) = [\mathfrak{p}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_k]$ and if $\mathfrak{p}^{(i)}$ is the zero-dimensional prime ideal to which $\mathfrak{q}_i$ belongs, then $\mathfrak{q} \not\equiv 0(\mathfrak{p}^{(i)})$, $i = 1, 2, \cdots, k$. This follows immediately from (1). If $\mathfrak{q} = \mathfrak{I} \cdot (\omega_1, \cdots, \omega_k)$, then for non-special constants u_i, the quantity $u_1\omega_1 + \cdots + u_h\omega_h$ is not divisible by any one of the ideals $\mathfrak{p}^{(i)}$, $i = 1, 2, \cdots, k$. We are therefore justified in assuming that the base elements ω_i of $\mathfrak{q}$ satisfy the conditions $\omega_i \not\equiv 0(\mathfrak{p}^{(j)})$, $i = 1, 2, \cdots, h$; $j = 1, 2, \cdots, k$. If $F_i(\xi, \omega_i = \omega_i^n + a_{i1}\omega_i^{n-1} + \cdots + a_{in} = 0$ $(a_{ij} = a_{ij}(\xi_1, \cdots, \xi_r))$ is the equation which expresses the integral dependence of ω_i on $\xi_1, \cdots, \xi_r$, then the fact that $\omega_i \not\equiv 0(\mathfrak{p}^{(j)})$ implies that $\partial F_i/\partial\omega_i \not\equiv 0(\mathfrak{p}_0)$, (Zariski 4). If we write $F_i = \omega_i G_i(\xi_1, \cdots, \xi_r; \omega_i) + A_i(\xi_1, \cdots, \xi_r) = 0$, then $\partial F_i/\partial\omega_i = \omega_i(\partial G_i/\partial\omega_i) + G_i$. Since $\partial F_i/\partial\omega_i \not\equiv 0(\mathfrak{p}_0)$, we conclude that $G_i \not\equiv 0(\mathfrak{p}_0)$. Consequently, $G_i \not\equiv 0(\mathfrak{p}_s)$, and the quantity $\omega_i = -A_i/G_i$ is in $\mathfrak{N}$. On the other hand, $A_i \equiv 0(Q^*)$, where $Q^* = \mathfrak{q}^* \cap \mathfrak{M}$. Hence we have $\omega_i \equiv 0(\mathfrak{N} \cdot Q^*)$, $i = 1, 2, \cdots, h$. Since $\mathfrak{q}^* = \mathfrak{N} \cdot \mathfrak{q} = \mathfrak{N} \cdot (\omega_1, \cdots, \omega_h)$, we have $\mathfrak{q}^* \equiv 0(\mathfrak{N} \cdot Q^*)$, and hence $\mathfrak{q}^* = \mathfrak{N} \cdot Q^*$.

The converse will follow immediately from the part of the theorem already proved provided we prove the following

LEMMA. *If Q^* is any primary ideal in $\mathfrak{M}$ belonging to the prime P^*, then there exists a primary ideal $\mathfrak{q}^*$ belonging to $\mathfrak{p}^*$ in $\mathfrak{N}$ such that $Q^* = \mathfrak{q}^* \cap \mathfrak{M}$.*

Proof. If Λ is the quotient field of $\mathfrak{M}$, then Σ, the quotient field of $\mathfrak{N}$ is a finite algebraic extension of Λ. We let Ω denote the least normal extension of Λ which contains Σ. We assert that if $\bar{\mathfrak{o}}$ is the integral closure of $\mathfrak{M}$ in Ω, then the relation

$$\bar{\mathfrak{o}}Q^* \cap \mathfrak{M} = Q^* \tag{2}$$

holds for any primary ideal in $\mathfrak{M}$ with P^* as its associated prime. If η is any element of $\bar{\mathfrak{o}}Q^* \cap \mathfrak{M}$, we can write

$$\eta = \omega_1 q_1 + \cdots + \omega_g q_g, \quad \omega_i \subset \bar{\mathfrak{o}}, \quad q_i \equiv 0(Q^*).$$

If $\theta_1, \cdots, \theta_\mu$ are the elements of the Galois group $\mathfrak{g}(\Omega : \Lambda)$, then since η is in Λ we have

$$\mu\eta = \sum_{i=1}^{\mu} (\theta_i \times \eta) = (\Sigma\theta_i \times \omega_1)q_1 + \cdots + (\Sigma\theta_i \times \omega_g)q_g.$$

The quantities $\sum_i \theta_i \times \omega_j$ are elements of Λ and since $\mathfrak{M}$ is integrally closed in Λ they must be elements of $\mathfrak{M}$. Hence $\mu \cdot \eta \equiv 0(Q^*)$ and consequently $\eta \equiv 0(Q^*)$.

Any prime ideal $\bar{P}$ in $\bar{\mathfrak{o}}$ which is such that $\bar{P} \cap \mathfrak{M} = P^*$ must be maximal in $\bar{\mathfrak{o}}$ (Krull 2). Moreover, it is proved in Krull 2 that since $\mathfrak{M}$ is integrally closed in Λ any two prime ideals in $\bar{\mathfrak{o}}$ which lie over P^* are conjugate. Hence if

$$\bar{\mathfrak{o}}Q^* = [\mathfrak{q}_1, \cdots, \mathfrak{q}_k] \tag{3}$$

is the normal decomposition of $\bar{\mathfrak{o}}Q^*$, and if $\bar{\mathfrak{p}}_i$ is the prime ideal associated with $\mathfrak{q}_i$, then the ideals $\bar{\mathfrak{p}}_1, \cdots, \bar{\mathfrak{p}}_k$ form a complete set of conjugate maximal ideals lying over P^*.

Let θ_i be an element of $\mathfrak{Q}(\Omega:\Lambda)$ such that $\theta_i \times \bar{\mathfrak{p}}_1 = \bar{\mathfrak{p}}_i$. The ideal $\mathfrak{q}'_i = \theta_i \times \mathfrak{q}_1$ is a primary ideal belonging to $\bar{\mathfrak{p}}_i$. If $\mathfrak{q}_i$ were not divisible by $\mathfrak{q}'_i$, we would have $\mathfrak{q}_i \supset [\mathfrak{q}'_i, \mathfrak{q}_i] \supset \bar{\mathfrak{o}}Q^*$, $\mathfrak{q}_i \neq [\mathfrak{q}'_i, \mathfrak{q}_i]$. This contradicts the fact that (3) is the normal decomposition of $\bar{\mathfrak{o}}Q^*$. We thus have $\mathfrak{q}_i \equiv 0(\mathfrak{q}'_i)$. Similarly we see that $\mathfrak{q}_1 \equiv 0(\theta^{-1} \times \mathfrak{q}_i)$. We therefore have $\mathfrak{q}_1 = \theta^{-1} \times \mathfrak{q}'_i \supseteqq \theta^{-1} \times \mathfrak{q}_i \supseteqq \mathfrak{q}_1$. Accordingly, we conclude that any two of the ideals $\mathfrak{q}_1, \mathfrak{q}_2, \cdots, \mathfrak{q}_k$ are conjugate. This proves that $\mathfrak{q}_i \cap \mathfrak{M} = Q^*$, $i = 1, 2, \cdots, k$, for if $\mathfrak{q}_i \cap \mathfrak{M}$ were a proper divisor of Q^* then $[\mathfrak{q}_1, \cdots, \mathfrak{q}_k] \cap \mathfrak{M}$ would be a proper divisor of Q^* also. This is not the case.

We now put $\mathfrak{o}_1 = \mathfrak{M}[\xi_{r+1}, \cdots, \xi_n]$. It is clear that $\mathfrak{o}_1 \subset \mathfrak{N}$ and $\mathfrak{o}_1 \subset \bar{\mathfrak{o}}$. One can see immediately that if $\mathfrak{p} = \mathfrak{p}^* \cap \mathfrak{o}_1$ then $\mathfrak{o}_{1\mathfrak{p}} = \mathfrak{N}$. Since $\bar{\mathfrak{o}}$ depends integrally on $\mathfrak{o}_1$ there is at least one prime in $\bar{\mathfrak{o}}$ which lies over $\mathfrak{p}$. This prime is obviously one of the ideals $\bar{\mathfrak{p}}_1, \cdots, \bar{\mathfrak{p}}_k$ say $\bar{\mathfrak{p}}_1$. The ideal $\mathfrak{q} = \mathfrak{q}_1 \cap \mathfrak{o}_1$ is primary, with $\mathfrak{p}$ as its associated prime. Moreover, $\mathfrak{q} \cap \mathfrak{M} = Q^*$. If then $\mathfrak{q}^* = \mathfrak{N} \cdot \mathfrak{q}$, $\mathfrak{q}^*$ is the desired ideal.

Theorem 7.2 enables us to extend the results on v-ideals which we have thus far obtained to the more general case of v-ideals of dimension $r-2$ in a ring of the type $\mathfrak{J} = K[\xi_1, \cdots, \xi_n]$, of degree of transcendency r over K, provided the v-ideals belong to valuations whose centers define simple subvarieties of the V_r associated with $\mathfrak{J}$. This is discussed in the next section.

8. V-ideals in rings of algebraic functions. We consider a field $\Sigma = K(\xi_1, \cdots, \xi_n)$ of algebraic functions of degree of transcendency r over K. The ring $\mathfrak{J} = K[\xi_1, \cdots, \xi_n]$ defines a model V_r of the field Σ. Let B be an $(r-2)$-dimensional valuation of Σ such that its valuation ring contains $\mathfrak{J}$. We assume that the center of B in $\mathfrak{J}$ is a prime $r-2$ dimensional ideal $\mathfrak{p}$ whose associated variety is a *simple* subvariety of V_r. We suppose that the non-homogeneous coordinates $\xi_1, \cdots, \xi_n$ of the general point of V_r have been selected so that if $R = K[\xi_1, \cdots, \xi_r]$, $P = \mathfrak{p} \cap R$, $\mathfrak{M} = R_P$ and $\mathfrak{N} = \mathfrak{J}_\mathfrak{p}$ then the relations between the primary $(r-2)$-dimensional ideals in $\mathfrak{N}$ and $\mathfrak{M}$ described in Theorem 7.2 hold. It is permissible to assume that $\xi_1, \cdots, \xi_{r-2}$ are algebraically independent modulo P. We denote the field $K(\xi_1, \cdots, \xi_{r-2})$ by Δ, and instead of ξ_{r-1} and ξ_r we write x any y. It is evident that the ring $\mathfrak{o} = \Delta[x, y]$ is between the rings $\mathfrak{M}$ and R, $R \subset \mathfrak{o} \subset \mathfrak{M}$. It is well known that the ideal $\bar{P} = \mathfrak{o}P$ is a prime ideal in $\mathfrak{o}$ of dimension zero with respect to the field Δ. Moreover, it is easily seen that the quotient ring $\mathfrak{o}_{\bar{P}}$ coincides with $\mathfrak{M}$.

Together with the valuation B we consider all $r-2$ dimensional valuations of Σ which have the ideal $\mathfrak{p}$ in $\mathfrak{J}$ as center. The ring $\mathfrak{R}$ is then contained in the valuation ring of any such valuation. Each valuation B of this set induces a valuation B_0 on the field $\Lambda = K(\xi_1, \cdots, \xi_r)$. Since $\xi_1, \cdots, \xi_{r-2}$ are algebraically independent modulo P, we conclude that $\mathfrak{o}$ is contained in the valuation ring $\mathfrak{B}_0$ of B_0. Moreover, the center of B_0 in $\mathfrak{o}$ is $\bar{P}$, and B_0 is of dimension zero with respect to the field Δ.

If $\mathfrak{q}$ *is an* $(r-2)$*-dimensional v-ideal in* $\mathfrak{R}$ *for the valuation* B, *then* $Q = \mathfrak{q} \cap \mathfrak{o}$ *is a zero-dimensional valuation ideal in* $\mathfrak{o}$ *for* B_0. *Moreover,* $\mathfrak{R} \cdot Q = \mathfrak{q}$. The fact that Q is a v-ideal for B_0 is obvious. To see that $\mathfrak{R} \cdot Q = \mathfrak{q}$ we need only note that since $\mathfrak{M} = \mathfrak{o}_{\bar{P}}$ we have $\mathfrak{M} \cdot Q = \mathfrak{q} \cap \mathfrak{M}$ and hence by Theorem 7.2, $\mathfrak{R} \cdot Q = \mathfrak{q}$.

It is a well known fact that if B_0 is any valuation of Λ which is of dimension zero with respect to the field Δ, and has its center at $\bar{P}$ in $\mathfrak{o}$, then there exists at least one valuation B of Σ whose valuation ring contains $\mathfrak{R}$ which induces the valuation B_0 on Λ. Moreover, this valuation B has its center at $\mathfrak{p}^*$ in $\mathfrak{R}$, and it is of dimension $r-2$. We conclude that *if* Q *is a zero-dimensional v-ideal in* $\mathfrak{o}$ *belonging to a valuation with origin at* $\bar{P}$, *then there is an* $r-2$ *dimensional valuation ideal* $\mathfrak{q}$ *in* $\mathfrak{R}$ *such that* $\mathfrak{q} \cap \mathfrak{o} = Q$. By our previous remark we must then have $\mathfrak{q} = \mathfrak{R} \cdot Q$.

We thus see that there is a $1:1$ correspondence between the $(r-2)$ dimensional v-ideals in $\mathfrak{R}$ and the zero-dimensional v-ideals in $\mathfrak{o}$ which belong to valuations with center at $\bar{P}$. The correspondence is such that if $\mathfrak{q}$ and Q are a corresponding pair, then $\mathfrak{q}$ is the extension of Q, Q the contraction of $\mathfrak{q}$. In view of this fact, the results which we have previously obtained for the zero-dimensional v-ideals in $\mathfrak{o}$ carry over automatically to the $r-2$ dimensional v-ideals in $\mathfrak{R}$. More precisely, all of the theorems of Section **5** are true if instead of reading "v-ideal in $\mathfrak{o}$" we read "v-ideal in $\mathfrak{R}$."

9. V-ideals and birational transformations. In order to expose the geometric significance of the simple v-ideals in $\mathfrak{R}$, it is necessary to have some type of birational transformation of V_r which will reduce a given simple ideal in kind. We can construct such transformations by slightly modifying the de Jonquières transformations studied in Section **6**.

As a base for the ideal P we have the quantities $f(x)$, $y-g(x)$ defined in Section **6**. The coefficients of f and g are in Δ. Since $\mathfrak{M} = \mathfrak{o}_{\bar{P}}$ we conclude that $P^* = \mathfrak{M} \cdot (f(x), y-g(x))$. We must then have $\mathfrak{p}^* = \mathfrak{R}(f(x), y-g(x))$. The functions $f(x)$ and $y-g(x)$ are, therefore, uniformizing parameters along the $r-2$ dimensional subvariety V_{r-2} of V_r defined by $\mathfrak{p}(\mathfrak{p} = \mathfrak{p}^* \cap \mathfrak{J})$.

If B is an $r-2$ dimensional valuation of Σ with center at $\mathfrak{p}^*$ in $\mathfrak{R}$, we may assume that $v_B(f(x)) \leqq v_B(y-g(x))$. (If the opposite inequality holds, we may reverse the roles of x and y). The quantity $y' = (y-g(x))/f(x)$ is then contained in the valuation ring $\mathfrak{B}$ of B. We consider the ring $\mathfrak{J}' = \mathfrak{J}[y'] = K[\xi_1, \cdots, \xi_n, y']$. Since both $\mathfrak{J}$ and y' are in $\mathfrak{B}$ we have $\mathfrak{B} \supset \mathfrak{J}' \supset \mathfrak{J}$. Associated with $\mathfrak{J}'$ is the algebraic variety V'_r in S_{n+1} whose general point is $(\xi_1, \cdots, \xi_n, y')$. The varieties V'_r and V_r are birationally equivalent.

If $\mathfrak{A}$ is an ideal in $\mathfrak{J}$, and $\mathfrak{A}'$ an ideal in $\mathfrak{J}'$ we shall say that the subvarieties $V(\mathfrak{A}) \subset V_r$ and $V'(\mathfrak{A}') \subset V'_r$ correspond to one another in the birational correspondence between V_r and V'_r if $\mathfrak{A}' \cap \mathfrak{J} = \mathfrak{A}$. (This simple definition is made possible by the fact that the ring $\mathfrak{J}'$ contains $\mathfrak{J}$. A more general definition due to Zariski says that $V_\sigma \subset V_r$ and $V'_\rho \subset V'_r$ are corresponding irreducible subvarieties provided there exists a valuation of Σ which has center at V_σ on V_r and at V'_ρ on V'_r). Any prime ideal $\mathfrak{p}'$ in $\mathfrak{J}'$ such that $\mathfrak{p}' \cap \mathfrak{J} = \mathfrak{p}$ must obviously be of dimension greater than or equal to $r-2$. Moreover, for any such ideal $\mathfrak{p}'$ we have

$$\mathfrak{p}' \supset \mathfrak{J}'\mathfrak{p} \supset \mathfrak{J}'(f(x), y-g(x)) = \mathfrak{J}'f \cdot (1, y') = \mathfrak{J}'f.$$

It follows that every subvariety of V'_r which corresponds to the subvariety $V_{r-2}(\mathfrak{p})$ on V_r must be on the variety defined by $\mathfrak{J}'f$ on V'_r. In particular, the center $\mathfrak{p}'$ of the valuation B has the property $\mathfrak{p}' \cap \mathfrak{J} = \mathfrak{p}$, and $\mathfrak{p}'$ is of dimension $r-2$. Since this is true of the center in $\mathfrak{J}'$ of any $r-2$ dimensional valuation B (with center at $\mathfrak{p}$ in $\mathfrak{J}$) which satisfies the condition $v_B(f) \leqq v_B(y-g)$, it is obvious that there are infinitely many $r-2$ dimensional ideals $\mathfrak{p}'$ in $\mathfrak{J}'$ which contract to $\mathfrak{p}$. In this sense, the variety $V_{r-2}(\mathfrak{p})$ is fundamental in the correspondence between V_r and V'_r.

We fix our attention on a particular valuation B with the above property. We denote the quotient ring $\mathfrak{J}'_{\mathfrak{p}'}$ by $\mathfrak{R}'$. It is evident that $\mathfrak{B} \supset \mathfrak{R}' \supset \mathfrak{R}$, where $\mathfrak{B}$ is the valuation ring of B. We apply the de Jonquières transformation $J: x' = x, y' = (y-g(x))/f(x)$ to the ring $\mathfrak{o}$. This extends $\mathfrak{o}$ to the ring $\mathfrak{o}' = \Delta[x', y']$. If B_0 is the valuation of Λ which is induced by B, and if $\bar{P}'$ is the origin of B_0 in $\mathfrak{o}'$, then we denote the quotient ring $\mathfrak{o}'_{\bar{P}'}$ by $\mathfrak{M}'$. It is at once clear that $\mathfrak{M} \subset \mathfrak{M}' \subset \mathfrak{R}'$. The relation between the v-ideals in $\mathfrak{M}$ and $\mathfrak{M}'$ are the same as those which hold between the v-ideals in $\mathfrak{o}$ and those in $\mathfrak{o}'$ which were described in Section **6**.

THEOREM 9.1. *If $\mathfrak{q}'$ is an $r-2$ dimensional v-ideal in $\mathfrak{R}'$, then $\mathfrak{q}' \cap \mathfrak{M}' = Q'$ is a zero-dimensional v-ideal in $\mathfrak{M}'$, and $\mathfrak{R}'Q' = \mathfrak{q}'$. Conversely, if Q' is a zero-dimensional v-ideal in $\mathfrak{M}'$, then $\mathfrak{R}'Q' = \mathfrak{q}'$ is an $(r-2)$-dimensional v-ideal in $\mathfrak{R}'$, and $\mathfrak{q}' \cap \mathfrak{M}' = Q'$.*

Proof. The fact that $\mathfrak{q}' \cap \mathfrak{M}'$ is a v-ideal Q' in $\mathfrak{M}'$ is evident. To prove that $\mathfrak{N}'Q' = \mathfrak{q}'$ we need only show that $\mathfrak{q}' \equiv 0(\mathfrak{N}'Q')$. Let h be an integer such that $f^hQ' \cap \mathfrak{M}$ is a v-ideal Q in $\mathfrak{M}$. (By Theorem 6. 4 there exists an integer h such that $f^hQ' \cap \mathfrak{M} = Q$, and $\mathfrak{M}'Q = f^hQ'$.) We suppose that h is the smallest integer with this property. It is easily verified that $f^h\mathfrak{q}' \cap \mathfrak{M}' = f^hQ'$. As a consequence of this, $f^h\mathfrak{q}' \cap \mathfrak{M} = Q$. If $\mathfrak{q}$ denotes the ideal $f^h\mathfrak{q}' \cap \mathfrak{N}$, then $\mathfrak{q} \cap \mathfrak{M} = Q$. The ideal $\mathfrak{q}$ must divide $\mathfrak{N}Q$. Since $\mathfrak{N} \cdot Q$ is a primary ideal belonging to $\mathfrak{p}^*$, the congruence $\mathfrak{N}Q \equiv 0(\mathfrak{q})$ implies that $\mathfrak{q}$ is a primary ideal belonging to $\mathfrak{p}^*$. By Theorem 7. 2, we have $\mathfrak{N} \cdot Q = \mathfrak{q}$. We note that if $k \geqq h$, then $f^kQ' \cap \mathfrak{M} = P^{*k-h}Q$. Since $f^k\mathfrak{q}' \cap \mathfrak{M} = f^kQ'$ we conclude that $f^k\mathfrak{q}' \cap \mathfrak{M} = P^{*k-h}Q$. Hence if $\mathfrak{A} = f^k\mathfrak{q}' \cap \mathfrak{N}$, then $\mathfrak{A} \cap \mathfrak{M} = P^{*k-h}Q$ and $\mathfrak{A} = \mathfrak{N}(P^{*k-h} \cdot Q) = \mathfrak{p}^{*k-h} \cdot \mathfrak{q}$.

Let a be any element in $\mathfrak{q}'$. Since a is in $\mathfrak{N}'$, we can write $a = \phi/\psi$, where ϕ and ψ are in $\mathfrak{J}'$, and $\psi \not\equiv 0(\mathfrak{p}')$. Now as an element of $\mathfrak{J}'$, we can write

$$\phi = a_0 + a_1y' + \cdot \cdot \cdot + a_ny'^m \qquad a_i \subset \mathfrak{J} \subset \mathfrak{N}$$

Consequently,

$$\psi a = [a_0f^m + a_1f^{m-1}(y - g) + \cdot \cdot \cdot + a_n(y - g)^m]/f^n.$$

If σ is the larger of the two integers h and m, then we have $f^\sigma\psi a \equiv 0(f^\sigma\mathfrak{q}' \cap \mathfrak{N})$. Hence $f^\sigma\psi a \equiv 0(\mathfrak{p}^{*\sigma-h} \cdot \mathfrak{q})$. We thus see that

$$f^\sigma\psi a = p_1q_1 + \cdot \cdot \cdot + p_tq_t$$

where $q_1, \cdot \cdot \cdot, q_t$ are elements of Q, and $p_1, \cdot \cdot \cdot, p_t$ are in $\mathfrak{p}^{*\sigma-h}$. This follows from the fact that $\mathfrak{q} = \mathfrak{N} \cdot Q$. In $\mathfrak{M}'$, we have $q_i = f^hq'_i$, $q'_i \equiv 0(Q')$, and in $\mathfrak{N}'$, we have $p_i = f^{\sigma-h}n'_i$, where n'_i is an element of $\mathfrak{N}'$. Hence, in $\mathfrak{N}'$, we have

$$f^\sigma\psi a = f^\sigma(n'_1q'_1 + \cdot \cdot \cdot + n'_tq'_t).$$

We therefore conclude that $\psi a = \Sigma n'_iq'_i$ is in $\mathfrak{N}'Q'$. Since ψ is a unit in the quotient ring $\mathfrak{N}'$, $a \equiv 0(\mathfrak{N}'Q')$. This proves that $\mathfrak{N}'Q' = \mathfrak{q}'$.

Conversely, let Q' be a v-ideal in $\mathfrak{M}'$, and let B_0 be a valuation of Λ for which Q' is a v-ideal. Let B be an extension of B_0 to Σ whose valuation ring contains $\mathfrak{N}'$, and let $\mathfrak{q}'$ be that v-ideal for B in $\mathfrak{N}'$ which satisfies the condition $v_B(\mathfrak{q}') = v_B(\mathfrak{N}'Q')$. We must then have $\mathfrak{q}' \cap \mathfrak{M}' = Q'$. By the first part of the theorem this implies that $\mathfrak{N}'Q' = \mathfrak{q}'$.

COROLLARY. *The subvariety $V'(\mathfrak{q}')$ of V'_r which is associated with $\mathfrak{p}'$ is a simple $r - 2$ dimensional subvariety of V_r.*

Proof. Our theorem implies in particular that $\mathfrak{p}^{*\prime}$, the origin of B in $\mathfrak{N}'$ is the extended ideal of the ideal $P^{*\prime}$, the origin of B_0 in $\mathfrak{M}'$. Now $P^{*\prime} \cap \mathfrak{o}' = \bar{P}'$

is a prime zero dimensional ideal in $\mathfrak{o}' = \Delta[x', y']$. Just as the ideal P in $\mathfrak{o}$ has as basis the two elements $f(x)$ and $y - g(x)$ so also $\bar{P}'$ has a basis consisting of two elements. These two elements will serve as a basis for $\mathfrak{p}^{*\prime}$ in $\mathfrak{N}'$, since $\mathfrak{p}^{*\prime} = \mathfrak{N}'P'$. The two elements are therefore uniformizing parameters at $\mathfrak{p}^{*\prime}$. Hence $V'(\mathfrak{p}')$ is simple.

The question of the behavior of the v-ideals in $\mathfrak{N}$ under the birational transformation T between V_r and V'_r can now be quickly settled. In fact, it is essentially the same as the behavior of the v-ideals in $\mathfrak{M}$ under the transformation J.

THEOREM 9. 2. *If $\mathfrak{q}$ is an $r - 2$ dimensional v-ideal in $\mathfrak{N}$ for the valuation B, and if h is that integer such that $\mathfrak{q} \equiv 0(\mathfrak{p}^{*h})$, $\mathfrak{q} \not\equiv 0(\mathfrak{p}^{*h+1})$, then $\mathfrak{N}'\mathfrak{q} = f^h\mathfrak{q}'$ where $\mathfrak{q}'$ is a v-ideal for B in $\mathfrak{N}'$. Moreover, $f^h\mathfrak{q}' \cap \mathfrak{N} = \mathfrak{q}$.*

Proof. Let $Q = \mathfrak{q} \cap \mathfrak{M}$. Obviously, $Q \equiv 0(P^{*h})$, $Q \not\equiv 0(P^{*h+1})$. Hence $\mathfrak{M}'Q = f^hQ'$, where Q' is a v-ideal for the valuation B_0 which B induces on Λ. By Theorem 9. 1, $\mathfrak{N}'(f^hQ') = f^h\mathfrak{q}'$ where $\mathfrak{q}'$ is a v-ideal for B in $\mathfrak{N}'$. We thus have $\mathfrak{N}'Q = f^h\mathfrak{q}'$. Since $\mathfrak{q} = \mathfrak{N}Q$, we have $\mathfrak{N}'\mathfrak{q} = f^h\mathfrak{q}'$. Since $f^h\mathfrak{q}' \cap \mathfrak{M} = Q$, it follows immediately that $f^h\mathfrak{q}' \cap \mathfrak{N} = \mathfrak{q}$.

In exactly the same manner, one can prove the following two theorems.

THEOREM 9. 3. *If $\mathfrak{q}'$ is an $r - 2$ dimensional v-ideal for B in $\mathfrak{N}'$ then there exists an integer h such that $f^h\mathfrak{q}'$ is the extended ideal of a v-ideal $\mathfrak{q}$ for B in $\mathfrak{N}$, and $f^h\mathfrak{q}' \cap \mathfrak{N} = \mathfrak{q}$. If k is any integer larger than h, then $f^k\mathfrak{q}' \cap \mathfrak{N} = \mathfrak{p}^{*k-h} \cdot \mathfrak{q}$, and $\mathfrak{N}' \cdot (\mathfrak{p}^{*k-h} \cdot \mathfrak{q}) = f^k\mathfrak{q}'$.*

THEOREM 9. 4. *If $\mathfrak{P}_h$ is the simple ideal of kind h for B in $\mathfrak{N}$, then $\mathfrak{N}'\mathfrak{P}_h = f^\sigma\mathfrak{P}'_{h-1}$ where $\mathfrak{P}'_{h-1}$ is the simple v-ideal of kind $h - 1$ for B in $\mathfrak{N}'$.*

We say that if $\mathfrak{N}'\mathfrak{q} = f_h\mathfrak{q}'$ then $\mathfrak{q}'$ is the transform of $\mathfrak{q}$ by the birational transformation T between V_r and V'_r; in symbols $T(\mathfrak{q}) = \mathfrak{q}'$.

10. Geometric interpretation of the results. The results which we have thus far obtained run parallel to those obtained by Zariski in his paper "Polynomial ideals defined by infinitely near base points." As is to be expected, an entirely analogous geometric interpretation can be given to them.

Let $\mathfrak{N}$ denote the quotient ring along a *simple* $r - 2$ dimensional subvariety V of an algebraic variety V_r. Let $\mathfrak{P}_{k+1}$ be a simple v-ideal of kind $k + 1$ in $\mathfrak{J}$. We associate with $\mathfrak{P}_{k+1}$ an $r - 2$ dimensional subvariety V_{k+1} infinitely near V and in the k-th neighborhood of V. The base locus

$$B(\mathfrak{P}_{k+1}) = V_1^{s_1}V_2^{s_2} \cdots V_k^{s_k}V_{k+1}$$

of $\mathfrak{P}_{k+1}$, and the effective multiplicity s_i of $\mathfrak{P}_{k+1}$ at V_i are defined by induction.

We define $B(\mathfrak{P}_1)$ to be $V_1 = V$. By Theorem 5.6 the v-ideals which precede $\mathfrak{P}_{k+1}$ are uniquely determined by $\mathfrak{P}_{k+1}$. Let $\mathfrak{q}_\lambda$ be the v-ideal which is the immediate predecessor of $\mathfrak{P}_{k+1}$, and suppose that $\mathfrak{q}_\lambda = \mathfrak{P}_1^{a_1} \cdot \mathfrak{P}_2^{a_2} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$ is the unique product decomposition of $\mathfrak{q}_\lambda$. $B(\mathfrak{P}_{k+1})$ is defined by the relation

$$B(\mathfrak{P}_{k+1}) = [B(\mathfrak{P}_1)]^{a_1} \cdot [B(\mathfrak{P}_2)]^{a_2} \cdots [B(\mathfrak{P}_k)]^{a_k} V_{k+1}.$$

It is postulated that if

$$\begin{aligned} & B(\mathfrak{P}_{i+1}) = V_1^{\mu_1} V_2^{\mu_2} \cdots V_i^{\mu_i} V_{i+1} \\ & B(\mathfrak{P}_{j+1}) = V_1^{\nu_1} V_2^{\nu_2} \cdots V_i^{\nu_j} V_{j+1} \\ \text{then:}\quad & B(\mathfrak{P}_{i+1}) \cdot B(\mathfrak{P}_{j+1}) = V_1^{\mu_1+\nu_1} \cdot V_2^{\mu_2+\nu_2} \cdots. \end{aligned}$$

If V_{r-1} is an $(r-1)$-dimensional subvariety of V_r which corresponds to an ideal $\mathfrak{A}$ in $\mathfrak{J}$, we shall say that V_{r-1} passes through the base locus $B(\mathfrak{P}_{k+1})$ with virtual multiplicities $s_1, s_2, \cdots, s_k, 1$ if

$$\mathfrak{N} \cdot \mathfrak{A} \equiv 0(\mathfrak{P}_{k+1}).$$

The following two properties of the base locus $B(\mathfrak{P}_{k+1})$ can be proved on the basis of our results in exactly the same way as the analogous properties are proved in Zariski 2 (Section 12).

1) *If* $B(\mathfrak{P}_{k+1}) = V_1^{s_1} \cdots V_k^{s_k} V_{k+1}$, *then* $\mathfrak{P}_{k+1} \equiv 0(\mathfrak{P}_1^{s_1})$, $\mathfrak{P}_{k+1} \not\equiv 0 (\mathfrak{P}_1^{s_1+1})$. Geometrically speaking, a V_{r-1} on V_r which passes through $B(\mathfrak{P}_{k+1})$ must have $V_1 = V$ as an s_1-fold subvariety.

2) *If* $\mathfrak{P}'_k = T(\mathfrak{P}_{k+1})$ *where* T *is the birational transformation studied in the preceding section, and if* $B(\mathfrak{P}_{k+1}) = V_1^{s_1} \cdots V_k^{s_k} V_{k+1}$ *then* $B(\mathfrak{P}'_k) = V'^{s_2}_1 \cdots V'^{s_k}_{k-1} V'_k$. Property 2) shows that the base locus $B(\mathfrak{P}_{k+1})$ behaves under the transformation T in exactly the same way as a set of points infinitely near a simple point of an algebraic surface behaves under a birational transformation which is locally quadratic.

The definition of the base locus can be extended to any power product of simple ideals by requiring that $B(\mathfrak{A} \cdot \mathfrak{B}) = B(\mathfrak{A}) \cdot B(\mathfrak{B})$ for any two ideal $\mathfrak{A}$ and $\mathfrak{B}$. If $\mathfrak{A}$ is a power product of simple v-ideals in $\mathfrak{N}$, and if A is any $r-1$ dimensional ideal in $\mathfrak{J}$, then the condition $\mathfrak{N} \cdot A \equiv 0(\mathfrak{A})$ is the arithmetic equivalent of the geometric condition that the variety $V(A)$ pass through a certain cluster of varieties infinitely near V with certain effective multiplicities.

As a final remark we point out that inasmuch as the transformation T reduces the kind of a simple ideal by one, it is always possible to "resolve" a given base locus by the successive application of a finite number of these transformations. In other words, after applying a finite succession of these transformations one obtains a variety V'_r birationally equivalent to V_r, on which all of the varieties in a given cluster of varieties infinitely near V_{r-2}'s appear as proper subvarieties at finite distance from one another.

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY

1. Krull, W., "Allgemeine Bewertungstheorie," *J. reine angew. math.*, vol. 167 (1932), pp. 160-196 (Krull 1).
2. Krull, W., "Zum Dimensionbegriff der Idealtheorie (Beiträge zur Arithmetik kommutativer Integritätsbereiche, III)," *Math. Zeit.*, vol. 42 (1937), pp. 745-766 (Krull 2).
3. Zariski, O., "Algebraic varieties over ground fields of characteristic zero," *American Journal of Mathematics*, vol. 62 (1940), pp. 187-221 (Zariski 1).
4. Zariski, O., "Polynomial ideals defined by infinitely near base points," *American Journal of Mathematics*, vol. 60 (1938), pp. 151-204 (Zariski 2).
5. Zariski, O., "The reduction of the singularities of an algebraic surface," *Annals of Mathematics*, vol. 40 (1939), pp. 639-689 (Zariski 3).
6. Zariski, O., "Some results in the arithmetic theory of algebraic varieties," *American Journal of Mathematics.*, vol. 61 (1939), pp. 249-294 (Zariski 4).

PROPERTIES OF THE COEFFICIENTS OF THE MODULAR INVARIANT $J(\tau)$.*

By D. H. Lehmer.

Introduction. The fundamental modular function $J(\tau)$, defined by (6), and known as Klein's absolute invariant, has been investigated for rather more than half a century. It is only in recent years, however, that some attention has been paid to the coefficients in the Fourier series for J. These coefficients become integers on multiplication by 12^3 and so it is customary to deal with the function

$$j(\tau) = 1728\,J(\tau)$$

whose Fourier development is

$$j(\tau) = e^{-2\pi i\tau} + 744 + 196884\,e^{2\pi i\tau} + 21493760\,e^{4\pi i\tau} + \cdots. \tag{1}$$

Introducing the variable x by

$$x = q^2 = e^{2\pi i\tau},$$

we may write (1) as [1]

$$j(\tau) = \sum_{k=-1}^{\infty} c(k)x^k.$$

A convergent series for $c(k)$ was discovered by Petersson [1] in 1932 and by Rademacher [1] in 1938. This series is too complicated to use in deriving the properties of $c(k)$ considered in this paper.[2] It shows at once, however, that

$$c(k) \sim 2^{-1/2}k^{-3/4}e^{4\pi\sqrt{k}} \tag{2}$$

so that $c(k)$ tends to infinity even more rapidly than the partition function $p(k)$.

The function $c(k)$ is, as a matter of fact, more closely related to Ramanujan's function $\tau(k)$ defined as the Fourier coefficient of the fundamental elliptic modular function

---

* Received July 26, 1941.

[1] It has been usual to write c_k for $c(k)$. The present notation avoids future typographical complications and is similar to that used for coefficients of other modular functions such as $\tau(k)$, $\sigma_3(k)$, $p(k)$ . . . .

[2] It was used by Rademacher [2], however, to show that a function $j(\tau)$ with these coefficients is indeed a modular invariant.

$$\Delta(1,\tau)(2\pi)^{-12} = h^{24}(\tau) = x\{(1-x)(1-x^2)(1-x^3)\cdots\}^{24} = \sum_{k=1}^{\infty}\tau(k)x^k.$$

$\tau(k)$ is of much lower order than $c(k)$. In fact Rankin [1] has proved that

$$(3) \qquad \tau(k) = O(k^{29/5}).$$

Perhaps the most interesting property of $\tau(k)$ is its multiplicative property

$$(4) \qquad \tau(m)\tau(n) = \tau(mn),$$

where m and n are coprime integers, discovered empirically by Ramanujan [1] and later proved by Mordell [1]. This enables one to compute $\tau(k)$ at once in terms of earlier τ's except when k is a power of a prime. In this case one has the formula

$$(5) \qquad \tau(p^a) = \tau(p)\tau(p^{a-1}) - p^{11}\tau(p^{a-2}),$$

so that the problem of determining $\tau(k)$ is reduced to the case in which k is a prime p. In Section **3** of the present paper we show that precisely the same is true of $c(k)$, namely that if k is composite, $c(k)$ may be expressed in terms of earlier c's alone. Such simple formulas as (4) and (5), which are compatible with (3), cannot be expected to hold for $c(k)$ in view of its asymptotic formula (2). When k is a prime, $c(k)$ is best calculated in terms of $\tau(n)$ and $\sigma_{11}(n)$ (the sum of the 11th powers of the divisors of n) by formula (10) derived in Section **1**.

In Section **2** part of the theory of the transformation equation of the n-th order has been modified to prepare the ground work for the properties of $c(k)$ established in Section **3**. In Section **4** certain formulas for the coefficients of $j^r(\tau)$ are derived, the multiplicative properties of which are briefly considered in Section **5**.

1. The function $c(k)$ and its congruence properties.

Klein's modular invariant $j(\tau)$ may be defined by any of the following expressions involving familiar elliptic functions:

$$\begin{aligned}(6) \qquad j(\tau) &= 12^3 J(\tau) = 2^8(1-k^2+k^4)^3 k^{-4}(1-k^2)^{-2} = [12g_2(1,\tau)]^3/\Delta(1,\tau)\\ &= [12g_2]^3/(g_2^3 - 27g_3^2) = 2^5[\theta_2^8+\theta_3^8+\theta_4^8]^3/\theta_2^8\theta_3^8\theta_4^8\\ &= (12Q)^3(Q^3-R^2)^{-1} = \{1+240\sum_{m=1}^{\infty} m^3x^m(1-x^m)^{-1}\}^3 x^{-1}\\ &\qquad\qquad \{(1-x)(1-x^2)(1-x^3)\cdots\}^{-24}\\ &= x^{-1} + 744 + 196884x + 21493760x^2 + 864299970x^3 + \cdots\\ &= \sum_{k=-1}^{\infty} c(k)x^k.\end{aligned}$$

The functions Q and R are, in the notation of Ramanujan:

$$(7) \qquad Q(x) = 1 + 240 \sum_{\nu=1}^{\infty} \sigma_3(\nu) x^{\nu}$$

$$(8) \qquad R(x) = 1 - 504 \sum_{\nu=1}^{\infty} \sigma_5(\nu) x^{\nu} = -504 \sum_{\nu=0}^{\infty} \sigma_5(\nu) x^{\nu}$$

in which $\sigma_r(n)$ is the sum of the r-th powers of the divisors of n, and

$$\sigma_5(0) = -1/504.$$

Comparing the 6-th definition of $j(\tau)$ with the 7-th we see that

$$Q^3 - R^2 = 12^3 x\{(1-x)(1-x^2)(1-x^3)\cdots\}^{24} = 1728 \sum_{m=1}^{\infty} \tau(m) x^m.$$

Hence we have

$$(9) \qquad (j(\tau) - 12^3) \sum_{m=1}^{\infty} \tau(m) x^m = R^2 = 504^2 \{ \sum_{\nu=0}^{\infty} \sigma_5(\nu) x^{\nu} \}^2$$
$$= 254016 \sum_{n=0}^{\infty} \Sigma_{55}^{(n)} x^n$$

where

$$\Sigma_{55}^{(n)} = \sigma_5(0)\sigma_5(n) + \sigma_5(1)\sigma_5(n-1) + \cdots + \sigma_5(n)\sigma_5(0).$$

This numerical function is not one of the 9 functions of Ramanujan which can be expressed in terms of $\sigma_r(k)$ alone. His analysis, however, may be applied with the result that $\Sigma_{55}^{(n)}$ involves both $\sigma_{11}(n)$ and $\tau(n)$ as follows:

$$\Sigma_{55}^{(n)} = \frac{65}{174132} \sigma_{11}(n) - \frac{3}{691} \tau(n).$$

Substituting this into (9) and identifying coefficients of x^m on both sides we have

$$(10) \quad \tau(m+1) + 24\tau(m) + c(1)\tau(m-1) + c(2)\tau(m-2) + \cdots + c(m$$
$$= \frac{65520}{691} (\sigma_{11}(m) - \tau(m))$$

a formula from which $c(k)$ may be found recursively in terms[3] of $\tau(n)$.

Ramanujan's remarkable congruence

$$\tau(m) \equiv \sigma_{11}(m) \pmod{691},$$

follows from this formula. If we were to take $c(0)$ to be 24 instead of 744, we could write (10) in the form[4]

---

[3] A table of $\tau(n)$ for $n \leq 30$ appears in Ramanujan [1]. A manuscript table of $\tau(n)$ for $n \leq 200$ is in the possession of the author.

[4] Another reason for setting $c(0) = 24$ is that this is the result of substituting $k = 0$ in the Petersson-Rademacher series for $c(k)$. The invariant properties of $J(\tau)$ are of course independent of $c(0)$ and our final formulas of Section **3** do not actually involve this additive constant.

$$c(k) = -\sum_{\nu=-1}^{k-1} c(\nu)\tau(k+1-\nu) + 65520(\sigma_{11}(k+1) - \tau(k+1))/691.$$

A table of $c(k)$ for $k \leq 24$ has been given by Zuckerman [1]. It was computed by a formula relating $c(k)$ to $p(25m-1)$ for $m \leq k$, values of the partition function being obtainable from tables of $p(k)$ for $k \leq 600$. This same formula can be inverted to compute values of $p(25k-1)$, beyond the range of tables, in terms of $c(m)$ for $m \leq k$. For example by (22) with $k = n = 5$, we readily find that

$$c(25) = 12\ 18832\ 84330\ 42251\ 04333\ 51500,$$

from which it follows that

$$p(624) = \quad 15289\ 37881\ 13516\ 82750\ 63375.$$

Zuckerman's values of $c(k)$ have been recomputed by methods of the present paper and found to be quite without error.

The functions $\tau(k)$ and $c(k)$ both possess congruence properties with respect to small moduli, a few of which we mention here.

The congruence

$$(k-1)\tau(k) \equiv 0 \qquad (\text{mod } 24)$$

is implied by one of Ramanujan's formulas for $\tau(k)$ and follows at once by taking the derivative of both sides of the generating identity

$$\{(1-x)(1-x^2)\cdots\}^{24} = \sum_{k=1}^{\infty} \tau(k)x^{k-1}.$$

An analogous congruence

$$(k+1)c(k) \equiv 0 \qquad (\text{mod } 24)$$

follows from the derivative of the identity

$$(11) \qquad xj(\tau) = \sum_{k=-1}^{\infty} c(k)x^{k+1} = Q^3[\sum_{\nu=0}^{\infty} p(\nu)x^\nu]^{24} - [\sum_{\nu=0}^{\infty} p(\nu)x^\nu]^{24} + 24I,$$

$p(\nu)$ being the partition function and I (here, and in what follows) denoting a power series with integral coefficients.

For the modulus 5 we have

$$(12) \qquad c(k-i) \equiv p(k/25)$$

$$-p\left(\frac{k-1}{25}\right) - p\left(\frac{k-2}{25}\right) + p\left(\frac{k-5}{25}\right) + p\left(\frac{k-7}{25}\right) - \cdots \quad (\text{mod } 5)$$

where the numbers $0, 1, 2, 5, 7, 12, \cdots$ are the pentagonal numbers $(3k^2 \pm k)/2$ and where $p(x) = 0$ if x is not an integer.

To prove this we may appeal to the theorem that if $\phi(x)$ is any power series with integer coefficients then

$$[\phi(x)]^{25} = \phi(x^{25}) + 5I.$$

Noting, from (7), that

$$Q = 1 + 5I$$

and writing (11) in the form

$$\begin{aligned}\sum_{k=0}^{\infty} c(k-1)x^k &= [\sum_{\nu=0}^{\infty} p(\nu)x^\nu]^{24} + 5I \\ &= [\sum_{\nu=0}^{\infty} p(\nu)x^\nu]^{25} \prod_{\lambda=1}^{\infty} (1-x^\lambda) + 5I \\ &= [\sum_{\nu=0}^{\infty} p(\nu)x^{25\nu}][\sum_{m=-\infty}^{\infty}(-1)^m x^{(3m^2+m)/2}] + 5I\end{aligned}$$

the congruence (12) follows from identifying coefficients of x^k. If k is not congruent to a pentagonal number modulo 25, that is, if $24k + 1$ is a non-residue of 25, then all the terms on the right of (12) are zero. But if $24k + 1$ is a non-residue of 25, so also is $k - 1$ and conversely. Changing $k - 1$ to k we may state that

$$(13) \qquad c(k) \equiv 0 \pmod 5$$

for $k \equiv 2, 3, 5, 7, 8, 10, 12, 13, 15, 17, 18, 20, 22, 23 \pmod{25}$. The number $c(k)$ is divisible by 5 for many other values of k also, as a glance at the table or at (12) will show. The first c not congruent to ± 1 or $0 \pmod 5$ is

$$c(39) \equiv 3 \pmod 5.$$

For the modulus 7 the congruence analogous to (12) is more complicated: If $k \neq 2$

$$(14) \qquad 2c(2k-4) - c(k-2) + c(k/2-1)$$
$$\equiv p\left(\frac{k}{49}\right) - p\left(\frac{k-1}{49}\right) - p\left(\frac{k-2}{49}\right) + p\left(\frac{k-5}{49}\right) + \cdots \pmod 7.$$

For proof we may write (9) in the form

$$x(j(\tau) - 12^3) = R^2[\sum_{\nu=0}^{\infty} p(\nu)x^\nu]^{24}.$$

If we square both sides and use

$$[\phi(x)]^{49} = \phi(x^{49}) + 7I$$

and

$$R = 1 + 7I,$$

in view of (8) we have

$$(15) \qquad x^2(j^2(\tau) + 2j(\tau) + 1) = (\sum_{\nu=0}^{\infty} p(\nu)x^{49\nu})(\sum_{m=-\infty}^{\infty} (-1)^m x^{(3m^2+m)/2}) + 7I.$$

If we borrow the result (29) of Section **4** we obtain

$$2c(2k-4)+c(k/2-1)+2(c(0)+1)\,c(k-2)$$

as the coefficient of x^k on the left of (15) for $k\neq 2$. This gives (14) on comparing this coefficient with the corresponding one on the right of (15) and noting that

$$2(c(0)+1)\equiv 1490\equiv -1 \pmod 7.$$

The analogue of (13) is

$$2c(2k)+c(k/2)\equiv c(k) \qquad \pmod 7 \quad (k\neq 0)$$

if k is a quadratic residue of 49.

More complicated congruence properties of $c(k)$ exist for higher moduli,[5] but these we omit.

2. The absolute invariant $F_n(\tau)$. Before deriving the multiplicative properties of $c(k)$ we develop very briefly some properties of $j(\tau)$ which are closely allied to its so called transformation equation (invariant equation)

$$\Phi_n(j(n\tau),j(\tau))=0$$

The present treatment is really part of a simplified theory of an equation of which Φ_n is an irreducible factor.

Let n be any positive integer. In terms of n we define a set Σ of $\sigma_1(n)$ triads of all positive integers

$$T=(d,h,\delta)$$

for which

$$d\delta=n \quad\text{and}\quad 0<h\leq\delta.$$

For each triad T of Σ we define an associate triad

$$T'=(d',h',\delta')$$

as follows:

d' is the greatest common divisor of δ and h, so that
$h=\eta d'$, where η is prime to δ/d',
$\delta'=n/d'$ and
$h'=d\eta'$, where η' is the least *positive* solution of the congruence

$$\eta\eta'=-1 \qquad (\operatorname{mod}\ \delta/d')$$

Since

$$0<h'=d\eta'\leq d\delta/d'=n/d'=\delta',$$

T' belongs to Σ.

---

[5] For example (10) gives a congruence property modulo 65520. In connection with the modulus 23 we note an error in Hardy [1], p. 165 equation (10.4.3): $\tau(23)\not\equiv 0$ (mod 23).

THEOREM 1. *The triad associated with T' is T itself.*

Proof. Let $T'' = (d'', h'', \delta'')$ be the associate of T', then d'' is the greatest common divisor of δ' and h'. But

$$\delta' = n/d' = (\delta/d')d$$

and $h' = d\eta'$ and since η' is prime to δ/d',

$$d'' = d.$$

Also

$$\delta'' = n/d'' = n/d = \delta.$$

It remains to show that $h'' = h$. Now $h'' = d'\eta''$ where

$$\eta''\eta' \equiv -1 \pmod{\delta'/d''}.$$

But

$$\delta'/d'' = \delta'/d = n/dd' = \delta/d'$$

so that

$$\eta'' \equiv \eta \pmod{\delta/d'}.$$

Therefore

$$h'' = d'\eta = h.$$

Hence the theorem is proved. We have at once

THEOREM 2. *The set of the associates of all the triads T of Σ is the set Σ itself.*

THEOREM 3. *Let $f(\tau)$ be any absolute modular invariant so that*

$$f(\tau) = f\left(\frac{a_1\tau + a_2}{a_3\tau + a_4}\right)$$

whenever $a_1a_4 - a_2a_3 = 1$. For each triad $T = (d, h, \delta)$ we define

$$f_T(\tau) = f((d\tau + h)/\delta).$$

Then

$$f_T(-1/\tau) = f_{T'}(\tau)$$

where T' is the associate of T.

Proof. Let

$$\tau_1 = \frac{d(-1/\tau) + h}{\delta} \qquad \tau_2 = \frac{d'\tau + h'}{\delta'}.$$

Then it may be verified that

$$\tau_2 = \frac{a_1\tau_1 + a_2}{a_3\tau_1 + a_4}$$

where

$$a_1 = \eta', \qquad a_2 = \frac{\eta\eta' + 1}{\delta/d'}, \qquad a_3 = \delta/d', \qquad a_4 = -\eta.$$

Moreover the a's are integers such that $a_1a_4 - a_2a_3 = 1$.
Hence

$$f(\tau_1) = f(\tau_2),$$

so that

$$f_T(-1/\tau) = f_{T'}(\tau).$$

THEOREM 4. *Let*

$$T_i = (d_i, h_i, \delta_i) \qquad (i = 1, 2, \cdots, \sigma_1(n))$$

be all the triads of Σ *and let*

$$f_{T_i}(\tau) = f((d_i\tau + h_i)/\delta_i).$$

Then the set of these $\sigma_1(n)$ *functions of* τ *is merely permuted when* τ *undergoes any transformation of the modular group.*

Proof. Since the group is generated by the two substitutions

$$\tau \to \tau + 1 \tag{16}$$

$$\tau \to -1/\tau \tag{17}$$

it suffices to prove the theorem for these substitutions alone. For (16) we have since $f(\tau + 1) = f(\tau)$,

$$f_{T_i}(\tau + 1) = f((d_i\tau + h + d_i)/\delta_i) = f_{T^*_i}(\tau)$$

where

$$T^*_i = (d_i, h^*_i, \delta_i),$$

the numbers h^*_i being the least positive residues of $h_i + d_i \pmod{\delta_i}$. It is clear that as T runs over Σ so also does T^*. For (17) we apply Theorem 3 and Theorem 2. This completes the proof of Theorem 4. It follows at once that any symmetric function of the $f_{T_i}(\tau)$ is an absolute modular invariant. In particular if the symmetric function is the sum and if $f(\tau) = j(\tau)$ we have:

THEOREM[6] 5. *The function*

---

[6] This theorem is a special case of Theorem 9 of Hecke [1] where the function j, which is of "type $(0, 1)$," is replaced by any modular form of type $(-k, 1)$ with $k \geq 0$. (His assumption that k be an even integer $\geqq 4$ may of course be dispensed with here). Theorem 5 also follows from the consideration of the roots of the equation

$$\prod_{\delta | n} \Phi_\delta(j(\delta\tau), j(\tau)) = 0$$

where $\Phi_\delta = 0$ is the classical transformation equation mentioned above. (See, for example, Fueter [1]).

$$(18)\qquad F_n(\tau) = \sum_T j_T(\tau) = \sum_{d\delta=n}\sum_{h=1}^{\delta} j((d\tau+h)/\delta).$$

is an absolute modular invariant.

THEOREM 6. *There exist $n+1$ integers*

$$A_0^{(n)}, A_1^{(n)}, \cdots, A_n^{(n)}, \qquad \textit{with } A_0^{(n)} = 1$$

such that

$$(19)\qquad F_n(\tau) = \sum_{\lambda=0}^{n} A_{n-2}^{(n)}\, j^{\lambda}(\tau)$$

Proof. If in (18) we expand each j in Fourier series we obtain

$$F_n(\tau) = \sum_{\delta|n}\sum_{m=-1}^{\infty} c(m)e^{2\pi i n\tau m/\delta^2}\sum_{h=1}^{\delta} e^{2\pi i h m/\delta}.$$

The sum over h has the value δ or zero according as m is or is not a multiple of δ. Setting $m = k\delta$, and $x = e^{2\pi i\tau}$, we obtain

$$(20)\qquad F_n(\tau) = x^{-n} + \sum_{\delta|n}\delta\sum_{k=0}^{\infty} c(k\delta)x^{kn/\delta}.$$

Since

$$j(\tau) = x^{-1} + 744 + 196884x + \cdots$$

it follows that $F_n(\tau) - j^n(\tau)$ has a pole of the $(n-1)$-st order at $x = 0$, and has integer coeficients. Hence there is an integer $A_1^{(n)}$ such that

$$F_n(\tau) - j^n(\tau) - A_1^{(n)}j^{n-1}(\tau)$$

has a pole of the $(n-2)$nd order at $x = 0$, and has integer coefficients. Finally there exists an integer $A_n^{(n)}$ such that

$$(21)\quad G(\tau) = F_n(\tau) - j^n(\tau) - A_1^{(n)}j^{n-1}(\tau) - \cdots - A_{n-1}^{(n)}j(\tau) - A_n^{(n)}$$

has a zero at $x = 0$. Now the function

$$G(\tau) = \sum_T j_T(\tau) - \sum_{\lambda=0}^{n} A_{n-\lambda}^{(n)}\, j^{\lambda}(\tau)$$

is regular and bounded in the fundamental triangle $(i, \omega, i\infty)$ of $j(\tau)$ wherever $j(\tau)$ itself is bounded, that is except perhaps at $\tau = i\infty$. But $G(\tau)$ is bounded there also, and in fact even tends to zero as $\tau \to i\infty$ since $x \to 0$ there. Finally $G(\tau)$ is an absolute invariant since $F_n(\tau)$ and $j^{\lambda}(\tau)$ are. Hence, by a well known theorem, $G(\tau)$ is a constant, and in fact $G(\tau) \equiv 0$, since $G(i\infty) = 0$. Hence the theorem follows from (21).

3. Multiplicative properties of $c(k)$. Our investigation of the coefficients $c(k)$ of $j(\tau)$ is closely connected with the coefficients of $F_n(\tau)$. These we denote by $s_n(k)$ so that

$$F_n(\tau) = \sum_{k=-n}^{\infty} s_n(k) x^k.$$

Inspection of (20) gives

THEOREM 7.

$$s_n(k) = \begin{cases} 1 & k = -n \\ 0 & -n \neq k < 0 \\ \sigma_1(n) c(0) & k = 0 \\ n \sum\limits_{\delta \mid (n,k)} c(kn/\delta^2)/\delta & k > 0 \end{cases}$$

where in the last case the sum extends over all divisors common to n and k

Thus

$$s_n(1) = nc(n)$$
$$s_p(k) = pc(pk) + c(k/p)$$

where p is a prime and $c(x) = 0$ if x is not an integer,

$$s_4(k) = 4c(4k) + [1 + (-1)^k] c(k) + c(k/4).$$

The following theorem expresses $s_n(k)$ in terms of $s_\nu(\lambda)$, with $\nu < n$, and is the source of our multiplication formulas.

THEOREM 8. *If $k \neq 0$, and $n > 1$, then*

$$s_n(k) = \sum_{\lambda=1}^{k+1} c(k-\lambda) s_{n-1}(\lambda) - \sum_{\mu=1}^{n-1} c(n-1-\mu) s_\mu(k) + c(n+k-1). \tag{22}$$

Proof. By Theorem 6

$$F_n(\tau) - j(\tau) F_{n-1}(\tau) = \sum_{\nu=0}^{n} \{A^{(n)}_{n-\nu} - A^{(n-1)}_{n-\nu}\} j^\nu(\tau)$$

where we take $A_n{}^{(n-1)}$ to be zero. This polynomial in $j(\tau)$ is really of the $(n-1)$-st degree and may be expanded as a linear combination of $F_\nu(\tau)$ for $\nu = 0, 1, 2, \cdots, n-1$, where, for sufficient reasons, we define

$$F_0(\tau) = -1.$$

Hence we write

$$F_n(\tau) - j(\tau) F_{n-1}(\tau) = \sum_{\nu=0}^{n-1} B^{(n-1)}_{n-1-\nu} F_\nu(\tau),$$

where the coefficients $B_0{}^{(n-1)}, \cdots, B^{(n-1)}_{n-1}$ are to be determined. If we imagine both sides expanded in powers of x, we find on equating coefficients of x^k for $k \neq 0$,

$$s_n(k) - \sum_{\lambda=1-n}^{k+1} c(k-\lambda) s_{n-1}(\lambda) = \sum_{\nu=1}^{n-1} B^{(n-1)}_{n-1-\nu} s_\nu(k). \tag{23}$$

Setting $k = -m$, where $1 < m < n$, and using Theorem 7 we find only a single term in each of the above sums different from zero, namely the first term of the left sum, and the m-th term of the right. Thus we have

$$-c(n-1-m) = B^{(n-1)}_{n-1-m} \qquad (m = 2, 3, \cdots, n-1).$$

For $k = -1$ we obtain

$$-c(n-2) - s_{n-1}(0) = B^{(n-1)}_{n-2}.$$

We have thus obtained all the B's that occur in (23). Substituting their values into (23) and recalling that $s_{n-1}(\lambda) = 0$ for $\lambda = -1, -2, \cdots, -n+2$ and that $s_{n-1}(1-n) = 1$, we obtain the formula (22).

It is worth pointing out that (22) does not actually involve $c(0)$. In fact $s_\lambda(0)$ does not occur and the two terms involving $c(0)$ cancel each other.

We are now in a position to consider multiplication theorems giving $c(nk)$ in terms of previous c's. These come from (22) on substituting for the s's in terms of the c's. The simplest of these is the duplication formula $(n = 2)$:

$$s_2(k) = \sum_{\lambda=1}^{k+1} c(k-\lambda)s_1(\lambda) - c(0)s_1(k) + c(k+1)$$

which becomes

$$2c(2k) = \sum_{\lambda=1}^{k-1} c(\lambda)c(k-\lambda) + 2c(k+1) - c(k/2). \tag{24}$$

For example for $k = 5$ and 6 we have

$$c(10) = c(6) + c(1)c(4) + c(2)c(3)$$
$$c(12) = c(7) + c(1)c(5) + c(2)c(4) + c(3)(c(3) - 1)/2.$$

Similarly for $n = 3$ the triplication formula is

$$\begin{aligned} 3c(3k) = \sum_{\lambda=1}^{k-1} c(k-\lambda)[2c(2\lambda + c(\lambda/2)] + 2c(2k+2) + c((k+1)/2) \\ + c(k+2) - c(k/3) - c(1)c(k), \end{aligned}$$

so that, for example,

$$\begin{aligned} 3c(12) = 2c(2)[c(4) + c(3)] \\ + c(1)[2c(6) - c(4) + c(2)] + c(6) + 2c(10). \end{aligned}$$

Quadruplication is accomplished by the formula

$$\begin{aligned} 4c(4k) = \sum_{\lambda=1}^{k-1} c(k-\lambda)[3c(3\lambda) + c(\lambda/3)] + 3c(3k+3) + c(k+3) \\ + c((k+1)/3) - c(k)[1 + (-1)^k] - c(1)[2c(2k) + c(k/2)] \\ - c(2)c(k) - c(k/4), \end{aligned}$$

and so on.

It should be noted that the first sum on the right of (22) involves the term $s_{n-1}(k+1)$ and hence $c(kn+n-k-1)$, so that (22) cannot be used to compute $c(nk)$ in terms of previous c's unless $n < k+1$. This condition can always be met by taking n as the smallest prime factor of nk. This also insures that the number $k+n+1$ of terms of (22) is a minimum. On the other hand (22) may be regarded as a formula for $c(kn+n-k-1) = c((k+1)(n-1))$. Quite a number of other expressions for $c(m)$ when m is composite[7] come out of (22). Thus if we note from Theorem 7 that

$$ks_n(k) = ns_k(n) \qquad (n > 0)$$

the equation (22) has, for example, the alternative form

$$s_n(k) = \sum_{\lambda=1}^{k+1} c(k-\lambda)s_{n-1}(\lambda) - 1/k \sum_{\mu=1}^{n-1} \mu c(n-1-\mu)s_k(\mu) + c(n+k-1).$$

4. Formulas for $A_\lambda^{(n)}$ and $c_\nu(k)$. Thus far we have used the mere existence of the integers $A_\lambda^{(n)}$ of Theorem 6 to derive our multiplicative properties. These numbers are of some interest since they give certain functional equations of $j(\tau)$, and are connected with the coefficients of $j^\nu(\tau)$. These latter we denote by $c_\nu(k)$ so that

$$j^\nu(\tau) = \sum_{k=-\nu}^{\infty} c_\nu(k)x^k.$$

Thus

$$c_\nu(-\nu) = 1, \quad c_\nu(1-\nu) = \nu c(0), \cdots, \qquad c_0(k) = \begin{cases} 1 \text{ if } k = 0 \\ 0 \text{ if } k \neq 0. \end{cases}$$

By (19) we have

$$(25) \qquad \sum_{k=-n}^{\infty} s_n(k)x^k = \sum_{\nu=0}^{n} \sum_{k=-\nu}^{\infty} A_{n-\nu}^{(n)} c_\nu(k)x^k.$$

Setting $k = 0, -1, -2, \cdots, -n$ and using Theorem 7 we have

$$\sum_{\nu=\mu}^{n} A_{n-\nu}^{(n)} c_\nu(-\mu) = \begin{cases} 1 & \text{if } \mu = n \\ 0 & \text{if } 0 < \mu < n \\ \sigma_1(n)c(0) & \text{if } \mu = 0. \end{cases}$$

These $n+1$ equations determine the $A^{(n)}$'s once the $c_\nu(-\mu)$ are known. These latter may be expressed in terms of $c(-1)$, $c(0)$, $c(1), \cdots$ (by means of the multinomial theorem) so that the same is true of the A's. Thus we find:

$$A_0^{(n)} = 1$$
$$A_1^{(n)} = -nc(0)$$
$$A_2^{(n)} = \binom{n}{2} c^2(0) - nc(1)$$

---

[7] When m is a prime, recourse may be had to the recurrence formula (10), as already mentioned.

$$A_3^{(n)} = -\binom{n}{3} c^3(0) + n(n-2)c(0)c(1) - nc(2)$$

$$A_4^{(n)} = \binom{n}{4} c^4(0) - n\binom{n-2}{2} c^2(0)c(1) + (n(n-3)/2)c^2(1)$$
$$+ n(n-3)c(0)c(2) - nc(3).$$

In these formulas for $A_\lambda^{(n)}$ it is understood that the superscript exceeds the subscript. For $A_n^{(n)}$ one must add $\sigma_1(n)c(0)$ to the above value. For example

$$A_4^{(4)} = c^4(0) - 4c^2(0)c(1) + 2c^2(1) + 4c(0)c(2) - 4c(3) + 7c(0)$$
$$= 85074\ 30000.$$

For $n = 2$, 3, and 4 equation (19) becomes

$$j(2\tau) + j(\tau/2) + j((\tau+1)/2) = j^2(\tau) - 1488\, j(\tau) + 162000$$

$$(26) \quad j(3\tau) + j(\tau/3) + j((\tau+1)/3) + j((\tau+2)/3)$$
$$= j^3(\tau) - 2232 j^2(\tau) + 1069956 j(\tau) - 368\ 64000$$

$$j(4\tau) + j(\tau+1/2) + j(\tau/4) + j((\tau+1)/4) + j((\tau+2)/4)$$
$$+ j((\tau+3)/4)$$
$$= j^4(\tau) - 2976 j^3(\tau) + 2533680 j^2(\tau) - 5614\ 44609 j(\tau)$$
$$+ 85074\ 30000.$$

The coefficients $c_\nu(k)$ of $j^\nu(\tau)$ are linearly expressible in terms of $s_\lambda(k)$ as we see at once from (25). The actual expression is

$$(27) \qquad c_\nu(k) = \sum_{\lambda=1}^{\nu} c_\nu(-\lambda) s_\nu(k) \qquad (k \neq 0).$$

In fact from (25) we have for $k \neq 0$

$$s_n(k) = \sum_{\nu=1}^{n} A_{n-\nu}^{(n)}\, c_\nu(k) \qquad (n = 1, 2, 3, \cdots, \nu).$$

These ν equations may be solved for the unknown $c_\nu(k)$ (since the determinant of the A's is equal to unity) as follows

$$(28) \qquad c_\nu(k) = \sum_{\lambda=1}^{\nu} a_\nu(\lambda) s_\lambda(k).$$

To determine the a's, which do not depend on k, we set $k = -\mu$, where $\mu = 1, 2, 3, \cdots, \nu$. By Theorem 7, all the terms of (28) are then zero, except that for which $\lambda = \mu$ so that we have

$$c_\nu(-\mu) = a_\nu(\mu) s_\mu(-\mu) = a_\nu(\mu).$$

This determination of the a's gives (27). Since $c_\nu(-\mu)$ is a polynomial in the numbers $c(0)$, $c(1)$, $c(2)$, $\cdots$; $c_\nu(k)$ may be expressed in terms of the c's as follows:

$$
\begin{aligned}
&c_2(k) = 2c(2k) + c(k/2) + 2c(0)c(k) \\
(29)\quad &c_3(k) = 3c(3k) + c(k/3) + 3c(0)[2c(2k) + c(k/2)] \\
&\qquad + 3[c^2(0) + c(1)]c(k) \\
&c_4(k) = 4c(4k) + c(k/4) + 4c(0)[3c(3k) + c(k/3)] \\
&\qquad + [6c^2(0) + 4c(1)][2c(2k) + c(k/2)] + [4c^3(0) \\
&\qquad + 12c(0)c(1) + 4c(2) + 1 + (-1)^k]c(k). \\
&\quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .
\end{aligned}
$$

In these formulas $k \neq 0$. For $k = 0$ additional terms are necessary; for example

$$c^3(0) + 3c(2) - 3c(0)c(1) - 4c(0)$$

in the case of $c_3(k)$.

5. Multiplicative properties of $c_\nu(k)$. The coefficients $c_\nu(k)$ of $j^\nu(\tau)$ possess multiplicative properties similar to, but more complicated than those of $c(k)$. These may be found in quite the same way as in Section **3** with the introduction of the function

$$F_n^{(\nu)}(\tau) = \sum_T j_T{}^\nu(\tau) = \sum_{\delta \mid n} \delta \sum_{k=-[\nu/\delta]}^{\infty} c_\nu(k\delta) x^{kn/\delta},$$

which is an absolute modular invariant, and hence a polynomial in $j(\tau)$ of degree $n\nu$. For example as analogue of (26)

$$
\begin{aligned}
F_2^{(2)}(\tau) = j^2(2\tau) + j^2(\tau/2) + j^2((\tau+1)/2) &= j^4(\tau) - 2976 j^3(\tau) \\
&+ 2535168 j^2(\tau) - 5636\;58750 j(\tau) + 87480\;00000.
\end{aligned}
$$

The counterpart of the duplication formula (24) is in this case

$$
\begin{aligned}
2c_2(2k) = \sum_{\lambda=-2}^{k+2} c_2(k-\lambda)[c_2(\lambda) - 2976c(\lambda)] + 2535168 c_2(k) \\
- 5636\;58750 c(k) - c_2(k/2) \qquad (k \neq 0).
\end{aligned}
$$

Substituting for the c_2's in terms of the c's from (29) gives incidentally another quadruplication formula for the c's, more complicated than the one found in Section **3**.

BIBLIOGRAPHY

W. E. H. Berwick:

[1]. "An invariant modular equation of the fifth order," *Quarterly Journal of Mathematics*, vol. 47 (1916), pp. 94-103.

R. Fueter:

[1]. *Vorlesungen über die singulären Moduln und die komplex Multiplication der elliptischen Funktionen.* Leipzig and Berlin, 1924. Part I, Chapter II.

G. H. Hardy:

[1]. *Ramanujan,* Cambridge, 1940 [Lectures IX, X, XII].

E. Hecke:

[1]. "Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Productentwicklung. I.," *Mathematischen Annalen,* vol. 114 ( ), pp. 1-28.

L. J. Mordell:

[1]. "On Mr. Ramanujan's empirical expansions of modular functions," Cambridge Philosophical Society, *Proceedings,* vol. 19 (1917), pp. 117-124. See also Hardy [1], pp. 161-163.

H. Petersson:

[1]. "Über die Entwicklungskoefficienten der automorphen Formen," *Acta Mathematica,* vol. 58 (1932), p. 202.

H. Rademacher:

[1]. "The Fourier coefficients of the modular invariant $J(\tau)$," *American Journal of Mathematics,* vol. 60 (1938), pp. 501-512.

[2]. "The Fourier series and the functional equation of the modular invariant $J(\tau)$," *American Journal of Mathematics,* vol. 61 (1939), pp. 237-248.

S. Ramanujan:

[1]. "On certain arithmetical functions," Cambridge Philosophical Society, *Transactions,* vol. 22 (1916), pp. 159-84. Reprinted in *Collected Papers of Srinivasa Ramanujan,* Cambridge, 1927, pp. 136-162. See also Hardy [1].

R. A. Rankin:

[1]. "Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions (II)," Cambridge Philosophical Society, *Proceedings,* vol. 35 (1939), pp. 357-372. See also Hardy [1], Lecture X.

H. S. Zuckerman:

[1]. "The computation of the smaller coefficients of $J(\tau)$," *Bulletin of the American Mathematical Society,* vol. 45 (1939), pp. 917-919.

QUATERNIONS AND SUMS OF THREE SQUARES.*

By Gordon Pall.

1. U. V. Linnik[1] has given essentially the following result:

Theorem 1. *Let p be an odd prime. Denote by $r(m)$ the number of pure and proper quaternions $x = i_1x_1 + i_2x_2 + i_3x_3$ of norm m, where m is a positive integer such that*

$$(1) \qquad m \neq 4k \text{ or } 8k + 7, \text{ and } (-m \mid p) = 1.$$

Let x_0 be a solution of the congruence $x_0^2 \equiv -m \pmod{p}$. For each x consider the right-divisor (unique up to a left unit factor) of $x_0 + x$:

$$(2) \qquad x_0 + x = zt, \; z \text{ and } t \text{ integral quaternions}, \; Nt = p.$$

Then, if m is sufficiently large, every quaternion t of norm p occurs among the $r(m)$ equations (2).

Linnik's proof is rather ingenious, but contains a serious error, in that on p. 377 he states that "the number of representations of a given binary quadratic form of determinant D as a sum of three squares does not exceed $c_\epsilon D^\epsilon$," and "this can be proved by methods similar to those of Gauss." This statement is false for forms of the type $kh^2(l\xi^2 + 2m\xi\eta + n\eta^2)$ if h is large (see our (41)); and Linnik applies[2] it for forms in which kh^2 may be as large as $\Delta^{1/2}$, $\Delta = ln - m^2$. Direct application, in his article, of the true result introduces a large factor which would seem to vitiate his proof.

In this article we shall revise his proof (which covers nineteen pages, and contains duplications, misprints, and superfluous details), and apply recent results[3] of our own to complete his demonstration of Theorem 1.

To facilitate comparison with Linnik's Russian paper we add the following remarks. His result, tantamount to Theorem 1, is stated on page 365. He does not formulate it as a separate theorem, but remarks that its proof is the

---

* Received June 4, 1941.

[1] "On the representation of large numbers by positive ternary quadratic forms," *Bull. of the Acad. of Sci. of the USSR, math. ser.*, vol. 4 (1940), pp. 363-402 (Russian).

[2] *Ibid.*, pp. 377 and 382.

[3] References will be made to Pall I and II: I. "On the arithmetic of quaternions," *Transactions of the American Mathematical Society*, vol. 47 (1940), pp. 487-500; II. "On the rational automorphs of $x_1^2 + x_2^2 + x_3^2$," *Annals of Mathematics*, vol. 41 (1940), pp. 754-766.

chief difficulty. The proposition which he applies our Theorem 1 to prove, and which is correct in view of our work, is as follows:

Let f be a positive ternary form with the invariants $\Omega = p$, $\Delta = 1$, where p is an odd prime. Let $(-f \mid p) = 1$. Then every large integer m prime to p and consistent with the generic conditions of f, is represented by f at least $c_1 h(-m)/(\log\log m \cdot \log\log\log m)$ times.

On pages 390-401 he somewhat sketchily extends this result (with m prime to 2Ω) to forms f of invariants $(\Omega, 1)$, where Ω is odd and contains at least one odd prime factor p such that $(-f \mid p) = 1$. His proof involves generalized quaternions, and contains the same errors as in the earlier case. The correction of these errors may be more difficult.

2. Notations. The letters $a, \cdots, e, t, \cdots, z$, and K, L denote integral quaternions of the type $a = a_0 + i_1 a_1 + i_2 a_2 + i_3 a_3$ with rational integers a_i; and $i_\alpha^2 = -1$, etc. Latin letters $f, \cdots, s$, and letters with subscripts (except the quaternion units i_α) denote rational integers. The letters $\kappa_1, \kappa_2, \cdots$ denote positive constants independent of m, and depending at most on p and ϵ. Here ϵ is any given positive number.

We call a *pure*, or a *vector*, if $a_0 = 0$; *proper* if $(a_0, a_1, a_2, a_3) = 1$, *proper* (mod k) if $(a_0, a_1, a_2, a_3, k) = 1$. The norm Σa_i^2 of a is written Na; the real part a_0, $\mathfrak{R}(a)$. Every a has eight left-associates $\pm a$, $\pm i_\alpha a$; we may speak of these as "one quaternion" instead of eight.

3. By adjusting unit factors we can confine t in (2) to $p + 1$ non-left-associate quaternions of norm p. Let us assume, for the sake of contradiction, that *one of these values t is missing among all $r(m)$ equations* (2). Then $\bar{t}$ (or its associate) is also missing. For (2) implies

$$x_0 + y = \bar{z}\bar{t}, \qquad y = -txt^{-1}, \tag{3}$$

where y is evidently pure, integral, and of norm m; also, y is proper, since a prime dividing y would divide m $(= Ny)$ and px $(= -\bar{t}yt)$.

Our assumption implies further that if $x_0^2 + m \equiv 0 \pmod{p^s}$, and we consider divisors v of norm p^s,

$$x_0 + x = uv, \qquad Nv = p^s, \tag{4}$$

and factor v as $\theta t' t'' \cdots t^{(s)}$ $(\theta = \pm 1 \text{ or } \pm i_\alpha)$, then neither t nor $\bar{t}$ occurs among these factors of norm p. For, from

$$x_0 + x = atb, \qquad Nb = p^r, \qquad r \geqq 0, \tag{5}$$

follows $x_0 + bxb^{-1} = bat$, where bxb^{-1} is another x.

Linnik has the happy idea of choosing s variable with m:

$$(6) \qquad m^{\frac{1}{2}+\tau} \leqq p^s < pm^{\frac{1}{2}+\tau},$$

where τ is the fixed fraction (between 0 and 1/5)

$$(7) \qquad \tau = -\tfrac{1}{2}\log(1-p^{-1})/\log p.$$

Hence $p^{-2\tau} = 1 - 1/p$, and we have

$$(8) \qquad p(p-1)^{s-1} < 2p^s(1-1/p)^s = 2(p^s)^{1-2\tau} < \kappa_1 m^{\frac{1}{2}-2\tau^2}.$$

Hence the number n of distinct divisors v which can occur in the $r(m)$ equations (4) cannot exceed $8\kappa_1 m^{\frac{1}{2}-2\tau^2}$. For, t' cannot be t (by our assumption), and t'' can be neither t nor $\bar{t}$ nor $\bar{t}'$, at least two of these being non-left-associate, and so on; (if $\bar{t}'$ and t'' were associates, v would be improper). We may observe that, s being variable, $p(p-1)^{s-1}$ is not of the same order of size as the full number $p^s + p^{s-1}$ of proper quaternions of norm p; this is the crucial point in Linnik's method.

Let these n distinct v's occur, respectively, for $a_1, a_2, \cdots, a_n$ distinct vectors x. Accordingly,

$$(9) \qquad a_1 + \cdots + a_n = 8r(m), \qquad n < 8\kappa_1 m^{\frac{1}{2}-2\tau^2}.$$

We shall use the result of C. L. Siegel[4] that

$$(10) \qquad \kappa_2 m^{\frac{1}{2}-\epsilon} < r(m) < \kappa_3 m^{\frac{1}{2}+\epsilon}.$$

Hence $a_1^2 + \cdots + a_n^2 \geqq (a_1 + \cdots + a_n)^2/n \geqq \kappa_4 m^{1-2\epsilon}/m^{\frac{1}{2}-2\tau^2}$, or

$$(11) \qquad a_1^2 + \cdots + a_n^2 \geqq \kappa_4 m^{\frac{1}{2}+2\tau^2-2\epsilon}.$$

We shall ultimately prove that (without any assumption)

$$(12) \qquad a_1^2 + \cdots + a_n^2 < \kappa_5 m^{\frac{1}{2}+\epsilon}$$

and shall thus obtain the desired contradiction.

We observe, as in connection with (3), that $\bar{v}$ occurs exactly as often as v. Let us call (x, y) a *conjugate pair* if x and y are proper vectors of norm m, and any right-divisors of norm p^s of $x_0 + x$ and of $x_0 + y$ are conjugates. The number of conjugate pairs is $a_1^2 + \cdots + a_n^2$.

4. Pairs (x, y) associated with binary quadratic forms. We formulate a result from a recent article[5] (Linnik uses a similar result due to Venkov).[6] Let $m > 1$, and let $[x]$ denote a set of four proper vectors

---

[4] "Über die Classenzahl quadratischer Zahlkörper," *Acta Arithmetica*, vol. 1 (1935), pp. 83-86.

[5] Pall I, pp. 495-497. The writer did not then know of Venkov's result.

[6] B. Venkov, "On the arithmetic of quaternions," *Bull. Acad. Sci. USSR*, VI series, vol. 16 (1922), pp. 205-240 (Russian).

$$x = i_1x_1 + i_2x_2 + i_3x_3, \qquad -i_1xi_1 = i_1x_1 - i_2x_2 - i_3x_3,$$
$$-i_2xi_2 = -i_1x_1 + i_2x_2 - i_3x_3, \qquad -i_3xi_3 = -i_1x_1 - i_2x_2 + i_3x_3, \tag{13}$$

of norm m. With every class of properly primitive binary quadratic forms ϕ of determinant m is associated a process, expressed by (15) and (16), whereby every $[x]$ is carried into a unique $[y]$, and no two distinct $[x]$'s go into the same $[y]$. For a fixed $[x]$, as ϕ ranges over the p. p. classes of determinant m, $[y]$ ranges over all proper y's of norm m such that

$$y \equiv x \pmod{2}, \qquad \text{if } m \equiv 1 \text{ or } 2 \pmod{4},$$
$$[y] \not\equiv [-x] \pmod{4}, \text{ if } m \equiv 3 \pmod{8}. \tag{14}$$

That is, y runs over $\frac{1}{3}$ or $\frac{1}{2}$ of all proper vectors of norm m.

If $\phi = [k, 2h, l]$, $h^2 + m = kl$, the process is defined by

$$h + x = KL, \qquad h + y = LK, \tag{15}$$

$$y = LxL^{-1} = K^{-1}xK, \tag{16}$$

where K and L are respectively of norms k and l.

5. Conditions for conjugate pairs associated with ϕ. Besides (15)-(16) if (x, y) is a conjugate pair associated with ϕ,

$$x_0 + x = uv, \qquad x_0 + y = u^*\bar{v}, \tag{17}$$

$$x_0^2 + m = qp^s, \qquad Nu = q = Nu^*, \tag{18}$$

$$Lx = yL, \qquad \bar{K}x = y\bar{K}. \tag{19}$$

If we replace x_0 by p^s, q becomes $q + 2x_0 + p^s$. We thus secure

$$(q, p) = 1. \tag{20}$$

We can take ϕ to be reduced, and hence

$$2\,|\,h\,| \leqq k \leqq l, \qquad kl \leqq 4m/3, \qquad k^2 \leqq 4m/3. \tag{21}$$

Lemma 1. *Let $m > m_1(p)$. For any conjugate pair associated with ϕ*

$$vK \textit{ is pure, and } p^s \mid \mathfrak{R}\,(v\bar{L}). \tag{22}$$

For, $L(x_0 + x) = (x_0 + y)L$, or $Luv = u^*\bar{v}L$. Since $(q, p) = 1$, $\bar{v}L$ ha the same right-divisors of norm p^s as $u^*\bar{v}L$.[7] We can set

$$\bar{v}L = wv, \qquad \bar{v}\bar{K} = w'v. \tag{23}$$

---

[7] Pall I, p. 488, Lemma 2.

Hence v is a right-divisor of the quaternion $z = \bar{v}L$, and of $\bar{z}$, and of $z + \bar{z} = 2z_0$. Hence $2z_0 = ev$, $p^s \mid 2z_0\bar{v}$ where $\bar{v}$ is proper, $p^s \mid z_0$:

$$p^s \mid \mathfrak{R}(\bar{v}L), \qquad p^s \mid \mathfrak{R}(\bar{v}\bar{K}). \tag{24}$$

Trivially, $\mathfrak{R}(\bar{v}\bar{K}) = \mathfrak{R}(vK) = \mathfrak{R}(Kv) = \mathfrak{R}(\bar{K}\bar{v})$. By (21_3) and (6),

$$|\mathfrak{R}(vK)| \leqq (Nv \cdot NK)^{\frac{1}{2}} = (p^s k)^{\frac{1}{2}} < k_6 m^{\frac{1}{2}+\frac{1}{2}\tau}. \tag{25}$$

Since $p^s \geqq m^{\frac{1}{2}+\tau}$, (22_1) follows. For $v\bar{L}$ we have similarly

$$|\mathfrak{R}(v\bar{L})| < p^{\frac{1}{2}} m^{\frac{1}{4}+\frac{1}{2}\tau} l^{\frac{1}{2}};$$

and would have $\mathfrak{R}(v\bar{L}) = 0$ if $l < 2m^{\frac{1}{2}+\tau-\eta}$ with $\eta > 0$. Hence:

LEMMA 2. *If for $\eta = \tau^2$, or for any fixed $\eta > 0$, we have*

$$k \geqq m^{\frac{1}{2}-\tau+\eta}, \tag{26}$$

then $v\bar{L}$ is pure for $m > m_2(p, \eta)$.

Since v and K are proper,[7a] we can set

$$\begin{aligned} &v = v't, \qquad K = \bar{t}K', \qquad Nt = p^\sigma, \qquad 0 \leqq \sigma \leqq s, \qquad Nv' = p^{s-\sigma}, \\ &NK' = k' = k/p^\sigma, \text{ while } v'K' \text{ is proper and pure.} \end{aligned} \tag{27}$$

The number of distinct values σ cannot exceed $\kappa_7 \log m < \kappa_8 m^\epsilon$.

From (15)-(19) we obtain

$$\begin{aligned} &x_0 + x = uv't, \qquad x_0 + y = u^*\bar{t}\bar{v}', \qquad h + x = \bar{t}K'L, \\ &h + y = L\bar{t}K', \qquad \bar{K}'tx = y\bar{K}'t. \end{aligned} \tag{28}$$

Hence $x' = txt^{-1}$ is a proper vector of norm m, and

$$x_0 + x' = tuv', \qquad K'y = x'K', \qquad h + x' = K'L\bar{t}. \tag{29}$$

Now $v'K' = -ed$, where $e = \bar{K}'$, $d = \bar{v}'$. Hence we have

$$-etued = etuv'K' = e(x_0 + x')K' = k'(x_0 + y), \text{ or} \tag{30}$$

$$k' \mid etued. \tag{31}$$

Since ed is proper, the greatest common left-divisor of K' and d is 1, and we can solve $K'z + dw = 1$ in integral z and w. Hence

$$k' \mid etuedw = etue(1 - K'z), \qquad k' \mid etue, \qquad \text{or}$$

$$\bar{K}'tu = z'K', \qquad z' \text{ integral.} \tag{32}$$

[7a] Pall I, p. 491, Theorem 5.

Since $\bar{K}'$ is proper, it follows as in (23) that $k' \mid \Re(etu)$, or

$$\bar{K}'tu = fk' + i_1b_1 + i_2b_2 + i_3b_3 = fk' + b, \tag{33}$$

where f, b_1, b_2, b_3 are rational integers. We also set

$$v'K' = i_1a_1 + i_2a_2 + i_3a_3 = a. \tag{34}$$

By (30) the real part of $(fk' + b)a = etuv'K'$ is $k'x_0$. Hence by forming the norm of $\bar{a}\xi + (fk' + b)\eta$ in two ways we obtain

$$\begin{aligned} k'(p^{s-\sigma}\xi^2 + 2x_0\xi\eta + p^\sigma q\eta^2) - (fk')^2\eta^2 \\ = (a_1\xi - b_1\eta)^2 + (a_2\xi - b_2\eta)^2 + (a_3\xi - b_3\eta)^2. \end{aligned} \tag{35}$$

Thus every conjugate pair associated with ϕ leads for some σ to a representation of $k'\psi$ as a sum of three squares, where

$$\psi = [p^{s-\sigma},\ 2x_0,\ p^\sigma q - f^2k']. \tag{36}$$

However, only those representations in which

(37) a is proper, and a, b have the same right-divisors of norm k',

need be considered in connection with conjugate pairs. We have

(38) the determinant of ψ is equal to $m - p^{s-\sigma}f^2k'$,

and since this cannot be negative, (6) shows that

$$|f| \leqq m^{\frac{1}{4}-\frac{1}{2}\tau}p^\sigma/k^{\frac{1}{2}}. \tag{39}$$

Conversely, no particular complex of values f, a, b can arise for a given σ from more than 64 conjugate pairs. For v' (and therefore K') has at most eight values as a divisor of norm $p^{s-\sigma}$ of a; tu is given by (33), and as $Nt = p^\sigma$ is prime to Nu, t has at most eight values; x' and y are determined by (29), and $x = t^{-1}x't$.

6. Forms ϕ with large minima.

We prove

LEMMA 3. *Let $m > m_2$. Let the minima k, k' of ϕ and ϕ' satisfy (26), and let (x, y) and (x, y') be conjugate pairs associated respectively with ϕ and ϕ'. Then $y = y'$.*

For by Lemmas 1 and 2 we can write (temporary notation)

$$\begin{aligned} Kv &= a = i_1a_1 + i_2a_2 + i_3a_3, & \bar{v}L &= b = i_1b_1 + i_2b_2 + i_3b_3, \\ K'v &= c = i_1c_1 + i_2c_2 + i_3c_3, & \bar{v}L' &= d = i_1d_1 + i_2d_2 + i_3d_3. \end{aligned}$$

The vector parts of ab and cd are equal:

$$Kv \cdot \bar{v}L = p^s(h + x), \qquad K'v \cdot \bar{v}L' = p^s(h' + x).$$

Hence, the first equality being determinant for determinant,

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix}, \text{ or } \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = 0.$$

Since $ab = p^s(h + x)$ is not real, a and b are not proportional. Hence there are real numbers λ and μ such that $d = \lambda a + \mu b$, $\bar{v}L' = \lambda Kv + \mu\bar{v}L = -\lambda\bar{v}\bar{K} + \mu\bar{v}L$, or $L' = -\lambda\bar{K} + \mu L$. By (19), $L'x = yL'$, $y = L'xL'^{-1} = y'$.

COROLLARY 1. *The number of conjugate pairs associated with all forms whose minima k satisfy* (26) *is, for large m, at most $r(m) < \kappa_3 m^{\frac{1}{2}+\epsilon}$.*

7. **The number of representations of a binary quadratic form as a sum of three squares.** Let N be the number of solutions of

$$(40) \qquad g(l\xi^2 + 2m\xi\eta + n\eta^2) = \sum_{i=1}^{3} (a_i\xi + b_i\eta)^2$$

in integers a_i, b_i. Here $\psi = [l, 2m, n]$ may be assumed to be positive, and properly or improperly primitive, $\Delta = ln - m^2$, and $g > 0$. There is a general formula of C. L. Siegel[8] for questions of this sort, by which it is necessary only to calculate explicitly the number of solutions of certain systems of congruences. This was done for (40) (among other examples) by Hel Braun[9] in the case $g^2\Delta$ odd, and our results agree with hers for that case. The general result for (40) is as follows:

$$(41) \qquad N = 24 \prod_p \chi(p), \text{ where}$$

$$\begin{aligned} \chi(p) &= \tfrac{1}{2}[1 + (-\sigma \mid E)(-\tau\alpha \mid k\psi)] \text{ if } p = 2, \\ &= (1 + \zeta^\delta\eta^\gamma)(p^{[(\gamma+1)/2]} - 1)/(p - 1) \\ &\quad + (2 - \sigma)(1 + \zeta + \zeta^2 + \cdots + \zeta^\delta)p^{\gamma/2} \text{ if } p > 2. \end{aligned}$$

Here the product is taken over all primes p; and for any p,

$g = p^\gamma k$, $\Delta = p^\delta E$, $\gamma, \delta \geqq 0$, $p \nmid kE$, $\sigma = \frac{1}{2}(3 - (-1)^\gamma)$,
$\tau = \frac{1}{2}(3 - (-1)^\delta)$; $\alpha = (-1 \mid E)$ if $p = 2$; $\zeta = (-k\psi \mid p)$ and
$\eta = ((-)^{\delta+1}E \mid p)$ if $p > 2$.

---

[8] "Über die analytische Theorie der Quadratischen Formen," *Annals of Mathematics*, vol. 36 (1935), pp. 527-606.

[9] "Über die Zerlegung quadratischer Formen in Quadrate," *Jour. für Math.*, vol. 178 (1937), pp. 34-64 (p. 62).

In the Legendre symbols ψ is to be replaced by any number represented by ψ and prime to p. Note that $\chi(p) = 1$ if $p \nmid 2g\Delta$.

In particular, if g is quadratfrei, so that γ is 0 or 1 for every p, $\chi(p) \leqq 1 + \delta$ if $\gamma = 0$, $\chi(p) \leqq 2$ if $\gamma = 1$, and $N/24$ does not exceed the number of divisors of $g\Delta$, which is $O(g\Delta)^{\epsilon}$.[10]

If we consider only the representations in which $(a_1, a_2, a_3, b_1, b_2, b_3) = 1$, N is not materially reduced; for $\chi(p)$ is replaced by

$$(1 + \zeta^{\delta}\eta^{\gamma})p^{[(\gamma-1)/2]} + (2 - \sigma)(1 + \zeta + \cdots + \zeta^{\delta})(p - 1)p^{\frac{1}{2}\gamma-1} \text{ when } \gamma \geqq 2.$$

8. Returning to (35) we write $k' = k''h^2$, where h^2 is the largest square in k', $h > 0$. Since $v'K'$ is proper and pure, $k' \not\equiv 0 \pmod 4$.

LEMMA 4. *Every representation* (35) *satisfying* (37) *is of the form*

$$a = zc\bar{z}, \qquad b = zd\bar{z}, \tag{42}$$

where z is a proper quaternion of norm h, and c and d satisfy

$$k''(p^{s-\sigma}\xi^2 + 2x_0\xi\eta + (p^{\sigma}q - f^2k')\eta^2) = \sum_{i=1}^{3}(c_i\xi - d_i\eta)^2. \tag{43}$$

By (37), $b\bar{a} \equiv 0 \pmod{k''h^2}$. Let z be the left-divisor of norm h (unique up to a right-unit factor) of the proper vector $a = zw = -\bar{w}\bar{z}$. Since $b\bar{a} \equiv 0 \pmod h$, Theorem 5′ of a recent article[11] shows that $b = zw'$. Then $ww' \equiv 0 \pmod h$, w is proper, and $h \mid Nw$. The right-divisor of w must be $\bar{z}$, and we have (42). By (42) and (35), $Nc = (Na)/h^2 = k''p^{s-\sigma}$, $Nd = (Nb)/h^2 = (k'qp^{\sigma} - (fk')^2)/h^2 = k''(qp^{\sigma} - f^2k')$, $-\Sigma c_id_i = \mathfrak{R}(cd) = \mathfrak{R}(zc\bar{z}zd\bar{z})/h^2 = \mathfrak{R}(ab)/h^2 = x_0k'/h^2 = x_0k''$.

Since d may be improper we write $k'' = k_1k_2$, $d = k_1d'$, where d' is proper $\pmod{k_2}$. We note that if $k'' \equiv 2 \pmod 4$, then $2 \mid \mathfrak{R}(cd)$, $4 \mid N(cd)$, hence $2 \mid cd$. We shall prove that (43) implies

$$dc = k''x_0 + k''y', \; y' \text{ pure and integral.} \tag{44}$$

For we have $k'' \mid Nc$, $k'' \mid Nd$, $k'' \mid \Sigma c_id_i$. Hence k_2 divides each of Nc, Nd', and $\Sigma c_id'_i$, where c and d' are proper. By Corollary 10,[12] k_2 divides the vector part of $d'c$; hence k'' divides the vector part of dc. By (43),

$$Ny' = m - p^{s-\sigma}f^2k''h^2. \tag{45}$$

---

[10] For example see Hardy and Wright, *Theory of Numbers*, Clarendon Press (1938), p. 259.

[11] Pall I, p. 492.

[12] Pall I, p. 493.

The condition $b\bar a \equiv 0 \pmod{k''h^2}$ requires that we find the number of proper quaternions z of norm h such that $zdc\bar z \equiv 0 \pmod{k''h}$, that is, by (44), such that $zy'\bar z \equiv 0 \pmod{h}$. Write $y' = h_1 y''$, $h = h_1 h_2$, where y'' is proper $\pmod{h_2}$. We have to solve

$$zy''\bar z \equiv 0 \pmod{h_2}, \quad Nz = h, \quad z \text{ proper}. \tag{46}$$

By (45) we see that $h_1^2 \mid m$ and that $Ny'' \equiv m/h_1^2 \pmod{h_2}$.

9. On the congruence (46). We recently proved[13]

Lemma 5. *Let x be a proper vector of norm n, h be odd and positive. The proper quaternions t of norm h such that $tx\bar t \equiv 0 \pmod{h}$ are the same as the right-divisors of norm h of $x_0 + x$, where x_0 ranges over the solutions $x_0 \pmod{h}$ of $x_0^2 \equiv -n \pmod{h}$.*

Let π be an odd prime. Varying our notations somewhat, we denote by t_r in this section *any* proper quaternion of norm π^r. In Lemmas 6 and 7, x is pure, and proper $\pmod{\pi}$. We have first

Lemma 6. *If $\pi \mid Nx$ and $t_r x \bar t_r \equiv 0 \pmod{\pi}$, where $r \geqq 1$, then the right-divisor of norm π of t_r is the right-divisor of norm π of x.*

For we can write $t_r = t_k t_1 t_q$, where $r = k + 1 + q$, $t_q x \bar t_q \not\equiv 0$, $t_1 t_q x \bar t_q \bar t_1 \equiv 0 \pmod{\pi}$. By Lemma 5, t_1 is the right-divisor of $t_q x \bar t_q$ of norm π. This contradicts the properness of $\bar t_q \bar t_1$ unless $q = 0$.

Let ν_s be the number of solutions $x_0 \pmod{\pi^s}$ of $x_0^2 \equiv -n \pmod{\pi^s}$, where $n = Nx$. The ν_s non-left-associate solutions t_s of

$$t_s x \bar t_s \equiv 0 \pmod{\pi^s} \tag{47}$$

will be denoted specifically by u_s. We have

Lemma 7. *Let $r \geqq s \geqq 0$, $r > 0$. The general solution t_r of*

$$t_r x \bar t_r \equiv 0 \pmod{\pi^s} \tag{48}$$

is given by $t_{r-s}u_s$ with t_{r-s} restricted only by the properness of $t_{r-s}u_s$. Hence the number of non-left-associate t_r satisfying (48) *is $\pi^{r-s}\nu_s$ if $s > 0$, $\pi^{r-1} + \pi^r$ if $s = 0$.*

The case $s = 0$ is trivial. Proceeding by induction assume the lemma to be true with $s - k$ $(1 \leqq k \leqq s)$ in place of s. We can factor any t_r as $t_{r-s}t_s$. If (48) holds, either i) $t_s x \bar t_s \equiv 0 \pmod{\pi^s}$ and t_{r-s} can be taken arbitrary, which is what we wish to prove, or ii) $v = (t_s x \bar{t}_s)/\pi^{s-k}$ is integral, and proper (mod π) for some k, $1 \leqq k \leqq s$. By assumption, in case ii), $t_s = t_k u_{s-k}$, whence $v = t_k w \bar{t}_k$, w integral. Now $t_{r-s} v \bar{t}_{r-s} \equiv 0$ (mod π^k, hence mod π), and by Lemma 6, the right-divisor of norm π of t_{r-s} is equal to that of v, hence to that of $\bar{t}_k$. This implies that t_r is improper, so that case ii) is impossible.

---

[13] Pall II, p. 763, Corollary 6.

Corollary 2. *Let π^q $(q \geqq 0)$ be the highest power of π dividing y'. The number of quaternions t_r such that $t_r y' \bar{t}_r \equiv 0$ (mod π^r) is*

$$\pi^{r-1}(\pi+1) \text{ if } q \geqq r \geqq 1, \quad \pi^q \nu \text{ if } q < r, \tag{49}$$

where ν is the number of solutions α of $\alpha^2 \equiv -N(y'/\pi^q)$ (mod π^{r-q}).

Corollary 3. *If y' is replaced by $\lambda a y' \bar{a}$ where $(\lambda N a, \pi) = 1$, the number of quaternions t_r in Corollary 2 is unchanged.*

Write $h = \pi_\mu^{r_\mu} \cdots \pi_2^{r_2} \pi_1^{r_1}$ in powers of distinct odd primes. Every proper z of norm h can be factored (uniquely up to left unit factors) as $z_\mu \cdots z_2 z_1$, with z_i of norm $\pi_i^{r_i}$. The congruence (46) becomes

$$z_\mu \cdots z_2 z_1 y' \bar{z}_1 \bar{z}_2 \cdots \bar{z}_\mu \equiv 0 \pmod{\pi_1^{r_1} \cdots \pi_\mu^{r_\mu}},$$

and reduces to the sequence of congruences

$$z_1 y' \bar{z}_1 \equiv 0 \pmod{\pi_1^{r_1}}, \quad z_2(z_1 y' \bar{z}_1)\bar{z}_2 \equiv 0 \pmod{\pi_2^{r_2}}, \cdots;$$

and the numbers of solutions are, by Corollary 3, respectively the same as those of $z_i y' \bar{z}_i \equiv 0$ (mod $\pi_i^{r_i}$), $(i = 1, 2, \cdots, \mu)$.

Corollary 4. *The number of solutions of (46) does not exceed $2^{\nu(h_1)} h_1 \xi$, where ξ is the number of solutions α of*

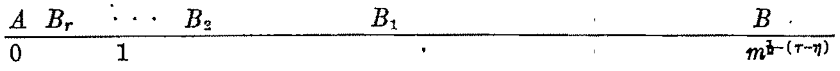
$$\alpha^2 \equiv -m/h_1^2 \pmod{h_2}. \tag{50}$$

Here $\nu(n)$ denotes the number of distinct prime factors of n, and we recall that $2^{\nu(n)} \leqq d(n) = O(n^\epsilon)$, where $d(n)$ is the number of divisors of n.

As for ξ, we have $\xi \leqq 2^{\nu(h_2)} h_3^{1/2}$, where $h_3 = (h_2, m/h_1^2)$. Hence the number of solutions of (46) does not exceed

$$(d(h))^2 \cdot (h_1^2 h_3)^{1/2} \leqq (d(h))^2 \rho^{1/2}, \text{ where } \rho = (m, k') = (m, k). \tag{51}$$

10. We count the conjugate pairs. Linnik classifies the values k not satisfying (26) into r intervals $B_i B_{i-1}$ where $r < \kappa_9 \log m < \kappa_{10} m^\epsilon$, as follows:



Here $AB_1 = m^{\frac{1}{2}-\nu_1} - \frac{1}{2}m^{\frac{1}{2}-(\tau-\eta)}$, $AB_2 = m^{\frac{1}{2}-\nu_2} = \frac{1}{2}m^{\frac{1}{2}-\nu_1}, \cdots, AB_r = m^{\frac{1}{2}-\nu_r} - \frac{1}{2}m^{\frac{1}{2}-\nu_{r-1}}$, $\frac{1}{2} \leqq m^{\frac{1}{2}-\nu_r} - AB_r < 1$, $\tau - \eta < \nu_1 < \nu_2 < \cdots < \nu_r$.

Suppose then that we have

$$(52) \qquad \tfrac{1}{2}m^{\frac{1}{2}-\nu} < k \leqq m^{\frac{1}{2}-\nu} (\leqq m^{\frac{1}{2}-(\tau-\eta)}).$$

Linnik states incorrectly that the number of reduced forms of determinant m with a given minimum k $(< m^{1/2})$ does not exceed $\kappa_{11} m^{\epsilon}$. The following example shows that this is false: $m = 2 \cdot 3^{4n}$, $k = 3^{2n}$, and the $2 \cdot 3^{n-1}$ forms $[3^{2n}, \pm 2 \cdot 3^n j, j^2 + 2 \cdot 3^{2n}]$ with $|j| \leqq 3^n/2$, $3 \nmid j$. The true result is that the number of such forms is equal to the number of solutions of $\alpha^2 \equiv -m \pmod{k}$, hence does not exceed $2^{\nu(k)}(k, m)^{1/2}$.

The number of sets of forms with k in (52) having a fixed value for $\rho = (k, m)$ is $d(m) < \kappa_{12} m^{\epsilon}$; and for a given σ (where $p^{\sigma} \mid k$), the number of values k with given ρ does not exceed $m^{\frac{1}{2}-\nu}/\rho p^{\sigma}$. The number of values of f is bounded in (39), where we can replace k by $\frac{1}{2}m^{\frac{1}{2}-\nu}$. Lastly, the number of values a, b does not exceed the number of values c, d satisfying (43) (which is $\leqq 24d(g_1) < \kappa_{13} m^{\epsilon}$, $g_1 = k''(m - p^{s-\sigma}f^2k')$) multiplied by the number of solutions z of (46), which we bounded in (51). Putting these together we see that the number of conjugate pairs associated with all forms not satisfying (26) does not exceed

$$\kappa_{14} \frac{m^{\frac{1}{2}-\nu}}{p^{\sigma}\rho} \rho^{\frac{1}{2}}\rho^{\frac{1}{2}} \frac{p^{\sigma} m^{1/4-\frac{1}{2}\tau}}{m^{1/4-\frac{1}{2}\nu}} (m^{\epsilon})^{7} = \kappa_{14} m^{\frac{1}{2}-\frac{1}{2}\nu-\frac{1}{2}\tau+7\epsilon}$$

for the least value of ν, which is $\tau - \eta$, hence $< \kappa_{15} m^{\frac{1}{2}-\tau+\frac{1}{2}\eta+\epsilon}$.

Comparing this with the result of § 6, we have (12).

McGill University,
Montreal.

ON BOUNDED ANALYTIC FUNCTIONS OF TWO COMPLEX VARIABLES IN CERTAIN DOMAINS WITH DISTINGUISHED BOUNDARY SURFACE.*

By Lipman Bers.

1. Introduction.

In order to generalize certain methods of the theory of analytic functions of one complex variable for the case of two variables, Bergman introduced the concept of domains with *distinguished boundary surface* (d. b. s). Such a domain is bounded by a finite number of *analytic hypersurfaces* (families of analytic surfaces depending on one real parameter). Their two dimensional intersection—the d. b. s. of the domain—plays a rôle analogous to that of the boundary curve in the theory of functions of one complex variable.[1]

The simplest example of a domain with d. b. s. is a product domain, say the unit bicylinder $|z_k| < 1$, $k = 1, 2$. It is bounded by the two analytic hypersurfaces $|z_k| = 1$, $|z_{3-k}| \leqq 1$; their intersection (the d. b. s.) is the surface $z_k = e^{i\theta_k}$.

A more general class of domains with d. b. s. (the only one we will consider in this paper) consists of domains $\mathfrak{M}$ [2] given by the relation:

$$z_1 \epsilon \mathfrak{B}(z_2), \ |z_2| < 1,$$

where $\mathfrak{B}(z_2)$ is the interior of a simple closed curve

$$\mathfrak{c}(z_2): \qquad z_1 = h(z_2, \lambda) \qquad (0 \leq \lambda < 2\pi)$$

and $h(z_2, \lambda)$ is an analytic function of z_2 for every fixed value of λ.[3] The bicylinder is also an $\mathfrak{M}$-domain. But it should be emphasized that an $\mathfrak{M}$-domain is in the general case not a pseudo-conformal transform of the bicylinder.[4]

---

* Received May 10, 1941; Revised October 28, 1941. This paper is a modified translation of a paper which was accepted by *Compositio Mathematica* but did not appear because of the war.

[1] See Bergman (1)-(8). The numbers in brackets refer to the bibliography.

[2] We denote manifolds by German letters, four dimensional ones by bold face capital letters, three dimensional by bold face small letters, two dimensional by capital letters and one dimensional by small letters. In operating with sets we use the usual symbols: ϵ (element), $\subset$ (subset), $\cdot$ (intersection), $+$ (sumset), $-$ (difference set), $E[\cdot\ \cdot\ \cdot]$ (set of points satisfying the relations enclosed in the brackets), $^{-}$ (closed set).

[3] Bergman (4), §§ 2, 3.

[4] A pseudo-conformal transformation is a one-to-one transformation by a pair of analytic functions of two complex variables.

In this paper we shall consider analytic and biharmonic functions[5] defined in $\mathfrak{M}$ and satisfying certain additional conditions (e. g. functions which are bounded or non-negative).[6] The theorem we shall prove can be understood as a generalization of some classic results of the theory of functions[7] (e. g. the Fatou theorem; a less strong extension of this theorem for the two variable case was previously obtained, by an entirely different method, by Bergman and Marcinkiewicz).[8]

2. The domain $\mathfrak{M}$ and its properties. Let $\mathfrak{B}(z_2)$ $(|z_2| \leq 1)$ be a star domain bounded by the curve

$$\mathfrak{c}(z_2): \qquad z_1 = h(z_2, \lambda), \qquad 0 \leq \lambda < 2\pi,$$

where h is a function satisfying the following conditions:

(a) h is a continuous function of its three real variables, an analytic function of z_2 for every fixed value of λ. One has

$$h(z_2,\lambda) - h(z_2,\lambda') \begin{cases} = 0 \text{ for } \lambda \equiv \lambda' & (\text{mod } 2\pi) \\ \neq 0 \text{ for } \lambda \not\equiv \lambda' & (\text{mod } 2\pi), \end{cases} \tag{2.1}$$

$$0 < 1/A < |h(z_2,\lambda)| < A. \tag{2.2}$$

(b) h possesses a continuous derivative of the second order with respect to the variable λ and

$$0 < 1/K < \left|\frac{\partial h(z_2,\lambda)}{\partial \lambda}\right| < K, \tag{2.3}$$

$$\left|\frac{\partial^2 h(z_2,\lambda)}{\partial \lambda}\right| < K' < \infty, \tag{2.4}$$

$$\frac{\partial \arg h(z_2,\lambda)}{\partial \lambda} > 1/B > 0. \tag{2.5}$$

---

[5] A function $u(x_1, y_1, x_2, y_2) = U(z_1, z_2)$ is called biharmonic if it satisfies the following system of differential equations:

$$\text{(i)} \qquad \frac{\partial^2 U}{\partial x_1^2} + \frac{\partial^2 U}{\partial y_1^2} = 0, \qquad \frac{\partial^2 U}{\partial x_2^2} + \frac{\partial^2 U}{\partial y_2^2} = 0,$$

$$\text{(ii)} \qquad \frac{\partial^2 U}{\partial x_1 \partial x_2} + \frac{\partial^2 U}{\partial y_1 \partial y_2} = 0, \qquad \frac{\partial^2 U}{\partial x_1 \partial y_2} - \frac{\partial^2 U}{\partial x_2 \partial y_1} = 0.$$

Real and imaginary parts of analytic functions of two complex variables are biharmonic functions and vice versa. A function satisfying the differential equations (i) is called doubly harmonic.

[6] See the preliminary report in Bers (1), (2).

[7] See Herglotz (1); Evans (1); Evans and Bray (1), (2); Riesz (1); Fatou (1) as well as the chapter on *beschränktartige* functions in Nevanlinna's *Eindeutige analytische Funktionen*, Berlin, 1936.

[8] See Bergman and Marcinkiewicz (1).

We define the domain $\mathfrak{M}$ as

$$\mathfrak{M} = E[z_1 = th(z_2, \lambda), \quad |z_2| < 1, \quad 0 \leq t < 1, \quad 0 \leq \lambda \leq 2\pi].$$

$\mathfrak{M}$ is bounded by the two analytic hypersurfaces

$$\bar{\mathfrak{e}}_1 = E[z_1 = h(z_2, \lambda), \quad |z_2| \leq 1, \quad 0 \leq \lambda \leq 2\pi]$$

and

$$\bar{\mathfrak{e}}_2 = E[z_1 = th(z_2, \lambda), \quad z_2 = e^{i\theta}, \quad 0 \leq t \leq 1, \quad 0 \leq \lambda \leq 2\pi \quad 0 \leq \theta \leq 2\pi]$$

and possesses the d. b. s.

$$\mathfrak{F} = E[z_1 = h(e^{i\theta}, \lambda), \quad z_2 = e^{i\theta}, \quad 0 \leq \theta < \pi, \quad 0 \leq \lambda < 2\pi].$$

To every Borel set $s \subset \mathfrak{F}$ corresponds a measurable set of "parameter points" $\{\theta, \lambda\}$. We shall call the (two dimensional) measure of this set the measure of s.

We shall use subdomains $\mathfrak{M}(r_1, r_2)$ $(0 < r_k < 1, k = 1, 2)$ of $\mathfrak{M}$ which are of the same type as $\mathfrak{M}$. They are defined as

$$\mathfrak{M}(r_1, r_2) = E[z_1 = th(z_2, \lambda), \quad |z_2| < r_2, \quad 0 < t < r_1, \quad 0 \leq \lambda < 2\pi].$$

Every point of $\mathfrak{M}$ is a point of the d. b. s. of some $\mathfrak{M}(r_1, r_2)$ and the sets $\mathfrak{M}(r_1, r_2)$ approach $\mathfrak{M}$ increasing as r_1, r_2 approach unity.

Let $w = W(z_1; z_2)$ be the function which maps $B(z_2)$ conformally into $|w| < 1$, $W(0; z_2) = 0$, $W[h(z_2, 0); z_2] = 1$. (W is an analytic function of z_1 but, in the general case, not an analytic function of z_2).

Lemma I. *For every fixed value of* z_2, W *possesses a continuous derivative* $W' = \partial W / \partial z_1$ *in* $\bar{\mathfrak{B}}(z_2)$ *and*

$$0 < 1/M < |W'| < M, \tag{2.6}$$

$$|W'(z_1; z_2) - W'(z'_1; z_2)| < N |z_1 - z'_1|. \tag{2.7}$$

We omit the simple proof which consists only of applying some theorems of Warschawski.[9]

Suppose

$$\begin{aligned} \rho(z_2; r, \lambda) &= |W[rh(z_2, \lambda); z_2]|, \\ T(z_2; r, \lambda) &= \arg W[rh(z_2, \lambda); z_2], \\ T(z_2; \lambda) &= T(z_2; 1, \lambda). \end{aligned}$$

For every fixed value of z_2 and r, T is a continuous increasing function of λ.

Lemma II. ρ *and* T *possses continuous derivatives* $\rho' = \partial \rho / \partial \lambda$ *and* $T' = \partial T / \partial \lambda$ *with respect to the variable* λ *and*

---

[9] Warschawski (1), p. 82.

$$r/AM < \rho(z_2; r, \lambda) < rAM, \tag{2.8}$$

$$\rho(z_2; r, \lambda) > 1 - AM(1 - r), \tag{2.9}$$

$$|\rho'(z_2; r, \lambda)| < \frac{D(1-r)}{r}, \tag{2.10}$$

$$1/KM < T'(z_2; \lambda) < KM, \tag{2.11}$$

$$|T'(z_2; r, \lambda) - T'(z_2; \lambda')| < \frac{C|\lambda - \lambda'| + D(1-r)}{r}, \tag{2.12}$$

$$\frac{1}{KM} - \frac{D(1-r)}{r} < T'(z_2; r, \lambda) < rAKM \tag{2.13}$$

(*C, D are constants*).

Proof. (2.8) and (2.9) follow from (2.6). Since

$$T(z_2; r, \lambda) = \frac{1}{i} \log \frac{W[rh(z_2, \lambda); z_2]}{\rho(z_2; r, \lambda)}$$

$$T'(z_2; r, \lambda) = \operatorname{Im} \frac{W'[rh(z_2, \lambda); z_2]}{W[rh(z_2, \lambda); z_2]} r \frac{\partial h(z_2, \lambda)}{\partial \lambda}. \tag{2.14}$$

On the other hand $\rho(z_2; 1, \lambda) = 1$; hence

$$T'(z_2; \lambda) = \frac{1}{i} \frac{W'[h(z_2, \lambda); z_2]}{W[h(z_2, \lambda); z_2]} \frac{\partial h(z_2, \lambda)}{\partial \lambda}, \tag{2.15}$$

whence (2.11) follows.

(2.12) is an immediate consequence of (2.14), (2.15), (2.3), (2.6), (2.7) and (2.8). By (2.12), (2.11) and (2.14) (2.13) is proved. We have

$$\rho'(z_2; r, \lambda) = \frac{1}{2} \frac{1}{|W|} \frac{\partial |W|^2}{\partial \lambda} = \operatorname{Re} \frac{W'[rh(z_2, \lambda); z_2]}{W[rh(z_2, \lambda); z_2]} r \frac{h(z_2, \lambda)}{\partial \lambda},$$
$$\rho'(z_2; 1, \lambda) = 0,$$

whence (2.10) follows (by (2.3), (2.6), (2.7) and (2.8)).

Suppose that

$$\{z_1, z_2\} \to \{h(e^{i\theta}, \lambda), e^{i\theta}\} \quad (\{z_1, z_2\} \in \mathfrak{M}). \tag{2.16}$$

We shall call this approach a *sectorial* one if

$$|z_1 - h(z_2, \lambda)| < a\,|z_1 - h(z_2, \lambda')| \quad (0 \leq \lambda' \leq 2\pi) \tag{2.17}$$

and

$$|z_2 - e^{i\theta}| < a(1 - |z_2|). \tag{2.18}$$

The approach

$$\{r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}\} \to \{h(e^{i\theta}, \lambda) e^{i\theta}\}$$

is a sectorial one. In fact, suppose that $\chi(z_2, \lambda)$ is the angle between the normal to the curve $\mathfrak{c}(z_2)$ at the point $h(z_2, \lambda)$ and the radius vector. We have

$$\operatorname{tg} \chi(z_2, \lambda) = \frac{1}{|h|} \frac{\partial |h|}{\partial \arg h}$$

so that, by (2. 2), (2. 3) and (2. 5),

$$|\chi(z_2, \lambda)| < \operatorname{arc\,tg} ABK < \frac{\pi}{2}$$

whence the existence of an a satisfying (2. 17) follows. Obviously (2. 9) is satisfied, too.

We shall say that a function U (defined in $\mathfrak{M}$) possesses the *sectorial limit* l at the point $\{h(e^{i\theta}, \lambda), e^{i\theta}\} \in \mathfrak{F}$ if $\lim U = l$ for every sectorial approach (2. 16).

3. **The function Q.** Let $G(z_2; z_1, z'_1)$ be the Green's function of the domain $\mathfrak{B}(z_2)$ and $n(z_2, \lambda)$ the direction of the inner normal to the curve $\mathfrak{c}(z_2)$ at the point $h(z_2, \lambda)$. Write

$$P(se^{it}, \phi) = \frac{1}{2\pi} \frac{1 - s^2}{1 - 2s\cos(t - \phi) + s^2},$$

$$\tilde{P}(z_2; z_1, \lambda) = \frac{1}{2\pi} \frac{\partial G[z_2; z_1, h(z_2, \lambda)]}{\partial n(z_2, \lambda)} \frac{\partial h(z_2, \lambda)}{\partial \lambda},$$

$$Q(z_1, z_2; \theta, \lambda) = \tilde{P}(z_2; z_1, \lambda) P(z_2, \theta),$$

$$Q(r_1, r_2; z_1, z_2; \theta, \lambda) = \tilde{P}(z_2; z_1/r_1, \lambda) P(z_2/r_2, \theta).$$

The function Q plays a part analogous to that of the kernal of the Poisson Integral; $Q(r_1, r_2; z_1, z_2; \theta, \lambda)$ is connected with $\mathfrak{M}(r_1, r_2)$ in the same way as $Q(z_1, z_2; \theta, \lambda)$ is with $\mathfrak{M}$.

Obviously,

$$\tilde{P}(z_2; z_1, \lambda) = P[W(z_1; z_2), T(z_2; \lambda)] T'(z_2, \lambda). \tag{3.1}$$

The following two theorems (proved by Bergman[10]) are an almost immediate consequence of the definition of Q.

A. *Suppose that U is biharmonic and regular in $\bar{\mathfrak{M}}$. Then*

$$U(z_1, z_2) = \int_0^{2\pi} \int_0^{2\pi} Q(z_1, z_2; \theta, \lambda) U[h(e^{i\theta}, \lambda), e^{i\theta}] d\theta d\lambda. \tag{3.2}$$

B. *Suppose that $\tilde{u}$ is a continuous function defined on $\mathfrak{F}$ and set $u(\theta, \lambda) = \tilde{u}[h(e^{i\theta}, \lambda), e^{i\theta}]$. Then*

---

[10] Bergman (5).

$$U(z_1, z_2) = \int_0^{2\pi}\int_0^{2\pi} Q(z_1, z_2; \theta, \lambda) u(\theta, \lambda)\, d\theta d\lambda \tag{3.3}$$

is continuous in $\bar{\mathfrak{M}}$ *and satisfies the following conditions:*

(i) $U(z_1, z_2^0)$ *is a harmonic function of* z_1,
(ii) $U[h(z_2, \lambda^0), z_2]$ *is a harmonic function of* z_2,
(iii) $U[h(e^{i\theta}, \lambda), e^{i\theta}] = u(\theta, \lambda)$.[11]

Analogous theorems are true for $M(r_1, r_2)$ and $Q(r_1, r_2; z_1, z_2; \theta, \lambda)$.

LEMMA III. *For every fixed point* $\{z_1, z_2\} \in M(R_1, R_2)$,

$$\lim_{r_1\to 1;\, r_2\to 1} Q(r_1, r_2; z_1, z_2; \theta, \lambda) = Q(z_1, z_2; \theta, \lambda)$$

(for $r_k > R_k$, $k = 1, 2$*) uniformly in* θ, λ.

We omit the very simple proof.

COROLLARY. *For sufficiently small* $1 - r_k$ $(k = 1, 2)$,

$$Q(r_1, r_2; z_1, z_2; \theta, \lambda) > q(z_1, z_2) > 0. \tag{3.4}$$

LEMMA IV. *Suppose* $\tau > 0$ *and*

$$\delta_{\Theta\Lambda}(\theta', \lambda') = \begin{cases} 1 \text{ for } \theta' \leq \Theta, \lambda' \leq \Lambda, \\ 0 \text{ for } \theta' > \Theta \text{ or } \lambda' > \Lambda. \end{cases} \tag{3.5}$$

For $|\theta' - \Theta| > \tau$, $|\lambda' - \Lambda| > \tau$

(3.6)

$$\left| \int_0^{\Theta}\int_0^{\Lambda} Q[r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}; \theta', \lambda']\, d\theta d\lambda - \delta_{\Theta\Lambda}(\delta', \lambda') \right| < \eta(\tau, 1 - r_1, 1 - r_2),$$

$$\lim_{\epsilon_1\to 0;\, \epsilon_2\to 0} \eta(\tau, \epsilon_1, \epsilon_2) = 0.$$

Proof. (3.6) is an immediate consequence of the following inequalities:

$$\int_0^{\Theta} P(re^{i\theta}, \theta')\, d\theta < \eta_1(\tau, 1 - r) \text{ for } 0 \leq \Theta \leq \theta' - \tau, \tag{3.7}$$

---

[11] Such functions form the "extended class of functions" (see Bergman (8)) and possess many properties in common with harmonic functions. In the case of a bicylinder the "extended class" is formed by the doubly harmonic functions (see footnote 5). The extended class of functions has many applications in the theory of analytic functions of two variables. It is possible to give another definition of the extended class which furnished also functions defined only in the interior of the domain. It will appear that every function which may be represented in the form (I) (see § 4) is a function of the extended class and that every non-negative function of the extended class can be represented in this form. We shall discuss this in another paper.

$$(3.8)\qquad 1-\int_0^{\Theta} P(re^{i\theta},\theta')\,d\theta < \eta_1(\tau,1-r) \text{ for } \theta'+\tau \leqq \Theta \leqq 2\pi,$$

$$(3.9)\qquad \int_0^{\Lambda} \bar{P}[z_2;rh(z_2,\lambda),\lambda']\,d\lambda < \eta_2(\tau,1-r) \text{ for } 0 \leqq \Lambda \leqq \lambda'-\tau,$$

$$(3.10)\qquad \left|\,1-\int_0^{\Lambda} \bar{P}[z_2;rh(z_2,\lambda),\lambda']\,d\lambda\,\right| < \eta_3(\tau,1-r) \text{ for } \lambda'+\tau \leqq \Lambda \leqq 2\pi$$

with

$$\lim_{\epsilon\to 0} \eta_k(\tau,\epsilon) = 0 \qquad (k = 1,2,3).$$

(3.7) and (3.8) follow from the obvious relations

$$P(r,t) < \frac{1}{2\pi}\cdot\frac{1-r}{r(1-\cos t)},\quad \int_0^{2\pi} P(re^{i\theta},\theta')\,d\theta = 1.$$

For $|\lambda-\lambda'| > \tau$ we have by (3.1)

$$\bar{P}[z_2;rh(z_2,\lambda),\lambda'] < \frac{1}{2\pi}\cdot\frac{1-\rho(z_2,r,\lambda)}{\rho(z_2,r,\lambda)[1-\cos(T(z_2,r,\lambda)-T(z_2,\lambda'))]}\,KM$$

so that, for sufficiently small $1-r$,

$$(3.11)\qquad \bar{P}[z_2;rh(z_2,\lambda),\lambda'] < \frac{1}{2\pi}\,\frac{A^2KM^3(1-r)}{r(1-\cos\tau/2KM)},$$

whence (3.9) follows. In order to prove (3.10), let us consider the integral

$$\begin{aligned} I(z_2,r,\lambda') &= \int_0^{2\pi} \bar{P}[z_2;rh(z_2,\lambda),\lambda']\,d\lambda \\ &= \int_0^{2\pi} P\{W[rh(z_2,\lambda);z_2],T(z_2,\lambda')\}T'(z_2,\lambda')\,d\lambda. \end{aligned}$$

From (2.9), (2.12) and (3.11) we have

$$\lim_{r\to 1}\,\{I(z_2,r,\lambda')-\int_0^{2\pi} P[\rho(z_2;r,\lambda),T(z_2;\lambda')-T(z_2;r,\lambda)]\,dT(z_2,r,\lambda)\} =$$

and from (2.10) we have

$$\lim_{r\to 1}\int_0^{2\pi} P[\rho(z_2,r,\lambda),T(z_2,\lambda')-T(z_2,r,\lambda)]\,dT(z_2,r,\lambda) = 1$$

uniformly in z_2 and λ'. Thus

$$(3.12)\qquad \lim_{r\to 1} I(z_2,r,\lambda') = 1$$

uniformly in z_2 and λ'. By (3.12) and (3.9) (3.10) is proved and the proof of the inequality (3.6) completed.

4. Integral representation of biharmonic functions. In this section we replace the Riemann Integral (3. 3) by a Stieltjes-Radon Integral over a completely additive (c. a.) set function $\omega(e)$ and look for biharmonic functions which may be represented in this form.[12] By e we denote here and hereafter a Borel set of points $\{\Theta, \Lambda\}$, Θ and Λ being understood as parameters on a torus surface (to which $\mathfrak{F}$ is topologically equivalent), say,

$$X = (2 + \cos\Theta)\cos\Lambda, \qquad Y = (2 + \cos\Theta)\sin\Lambda, \qquad Z = \sin\Theta$$

(X, Y, Z—Cartesian co-ordinates). ω is thus defined on a torus surface.

THEOREM. *Suppose that U is a biharmonic function defined in $\mathfrak{M}$,*

$$(4.1) \qquad 0 < r_k^\nu < 1, \ \lim_{\nu\to\infty} r_k^\nu = 1, \qquad k = 1, 2$$

and that

$$(4.2) \quad \int_0^{2\pi}\int_0^{2\pi} | U[r_1^\nu h(r_2^\nu e^{i\theta}, \lambda), r_2^\nu e^{i\theta}] \, | \, d\theta d\lambda < c < \infty \quad (\nu = 1, 2, \cdots).$$

Then U may be represented in the form

$$(\mathrm{I}) \qquad U(z_1, z_2) = \int_0^{2\pi}\int_0^{2\pi} Q(z_1, z_2; \theta, \lambda) d\omega(e_{\theta\lambda})$$

where ω is a c. a. set function.

COROLLARY. *If $U = V_1 - V_2$, V_k $(k = 1, 2)$ being biharmonic and non-negative, U may be represented in the form* (I).

Proof. Let us consider the c. a. set function

$$\omega_\nu(e) = \iint_e U[r_1^\nu h(r_2^\nu e^{i\theta}, \lambda), r_2^\nu e^{i\theta}] d\theta d\lambda.$$

By Theorem A (§ 3),

$$(4.3) \qquad U(z_1, z_2) = \int_0^{2\pi}\int_0^{2\pi} Q(r_1^\nu, r_2^\nu; z_1, z_2; \theta, \lambda) d\omega_\nu(e_{\theta\lambda}).$$

By (4. 2),

$$\int_0^{2\pi}\int_0^{2\pi} | \, d\omega_\nu \, | < c,$$

so that there exists a c. a. set function ω and a subsequence of $\{\omega_\nu\}$—for the sake of simplicity we call it $\{\omega_\nu\}$, too—such that

$$(4.4) \qquad \omega_\nu \to \omega \, (\nu \to \infty).^{13}$$

We have (for a fixed point $\{z_1, z_2\}$)

---

[12] Concerning the theorems of the theory of c. a. set functions, see Radon (1), (2).

[13] By " $\omega_\nu \to \omega$ " we denote weak convergence of the sequence of c. a. set functions ω_ν.

$$\left|\int_0^{2\pi}\int_0^{2\pi} Q(r_1^\nu, r_2^\nu)\,d\omega_\nu - \int_0^{2\pi}\int_0^{2\pi} Q\,d\omega\right| \leq$$
$$\int_0^{2\pi}\int_0^{2\pi} \mid Q(r_1, r_2) - Q \mid d\omega_\nu + \left|\int_0^{2\pi}\int_0^{2\pi} Q\,d(\omega - \omega_\nu)\right|.$$

The first term on the right hand is not greater than

$$c \max_{\substack{0\leq\theta\leq 2\pi\\ 0\leq\lambda\leq 2\pi}} \mid Q(r_1, r_2; z_1, z_2; \theta, \lambda) - Q(z_1, z_2; \theta, \lambda) \mid ,$$

hence it tends to 0 as $\nu \to \infty$ (see Lemma III). The same thing is true for the second term since Q is a continuous function of θ and λ. Thus

$$\lim_{\nu\to\infty} \int_0^{2\pi}\int_0^{2\pi} Q(r_1^\nu, r_2^\nu)\,d\omega_\nu = \int_0^{2\pi}\int_0^{2\pi} Q\,d\omega,$$

whence (I) follows by means of (4.3).

Proof of the Corollary. We may suppose that $V_2 = 0$, i. e. that $U > 0$. We have (see the corollary to Lemma III)

$$\int_0^{2\pi}\int_0^{2\pi} U[r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}]\,d\theta d\lambda$$
$$\leq \frac{1}{\min\limits_{\substack{0\leq\theta\leq 2\pi\\ 0\leq\lambda\leq 2\pi}} Q(r_1, r_2; z_1, z_2; \theta, \lambda)} \int_0^{2\pi}\int_0^{2\pi} Q(r_1, r_2)\,U\,d\theta d\lambda \leq \frac{U(z_1, z_2)}{q(z_1, z_2)}.$$

Let us now consider a function U which may be represented in the form (I), where ω is an *arbitrary* c. a. set function (see footnote 11).

THEOREM. *Suppose that* $U(z_1, z_2)$ $(\{z_1, z_2\} \in \mathfrak{M})$ *may be represented in the form* (I) *and set*

$$\omega_{r_1 r_2}(e) = \iint_e U[r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}]\,d\theta d\lambda \tag{4.5}$$

Then

$$\omega_{r_1 r_2} \to \omega \qquad (r_k \to 1, k = 1, 2). \tag{4.6}$$

Proof. We may suppose that $\omega \geq 0$ and that

$$\lim_{\epsilon\to 0}\left(\int_{-\pi}^{\pi}\int_{-\epsilon}^{\epsilon} d\omega + \int_{-\epsilon}^{\epsilon}\int_{-\pi}^{\pi} d\omega\right) = 0.$$

Set

$$\phi_{r_1 r_2}(\Theta, \Lambda) = \int_0^{\Theta}\int_0^{\Lambda} d\omega_{r_1 r_2} = \int_0^{\Theta}\int_0^{\Lambda} U[r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}]\,d\theta d\lambda,$$

$$\phi(\Theta, \Lambda) = \int_0^{\Theta}\int_0^{\Lambda} d\omega.$$

Then (4. 6) is equivalent to the assertion:

$$(4.7)\qquad \lim_{r_1\to 1;\, r_2\to 1} \phi_{r_1r_2}(\Theta, \Lambda) = \phi(\Theta, \Lambda)$$

for every point $\{\Theta, \Lambda\}$ *satisfying the conditions*

$$\lim_{\epsilon\to +0} \phi(\Theta - \epsilon, 2\pi) = \lim_{\epsilon\to +0} \phi(\Theta + \epsilon, 2\pi),$$

$$\lim_{\epsilon\to +0} \phi(2\pi, \Lambda - \epsilon) = \lim_{\epsilon\to +0} \phi(2\pi, \Lambda + \epsilon).$$

Suppose that these conditions are satisfied and that $\epsilon > 0$. We may choose a $\tau > 0$ such that

$$(4.8)\qquad \int_{\Theta-\tau}^{\Theta+\tau}\int_0^{2\pi} d\omega + \int_0^{2\pi}\int_{\Lambda-\tau}^{\Lambda+\tau} d\omega < \epsilon .$$

By (I),

$$\begin{aligned} &| \phi_{r_1r_2}(\Theta, \Lambda - \phi(\Theta, \Lambda) | \\ &= \left| \int_0^{\Theta} d\theta \int_0^{\Lambda} d\lambda \int_0^{2\pi}\int_0^{2\pi} Q[r_1h(r_2e^{i\theta}, \lambda), r_2e^{i\theta}; \theta', \lambda'] d\omega(e_{\theta'\lambda'}) - \int_0^{\Theta}\int_0^{\Lambda} d\omega \right| \\ &= \left| \int_0^{2\pi}\int_0^{2\pi} \left\{ \int_0^{\Theta}\int_0^{\Lambda} Q[r_1h(r_2e^{i\theta}, \lambda), r_2e^{i\theta}; \theta', \lambda'] d\theta d\lambda - \delta_{\Theta\Lambda}(\theta', \lambda') \right\} d\omega(e_{\theta'\lambda'}) \right| \\ &\leqq \int_0^{2\pi}\int_{\Lambda-\tau}^{\Lambda+\tau} + \int_{\Theta-\tau}^{\Theta+\tau}\int_0^{2\pi} + \int_0^{\Theta-\tau}\int_0^{\Lambda-\tau} + \int_0^{\Theta-\tau}\int_{\Lambda+\tau}^{2\pi} + \int_{\Theta+\tau}^{2\pi}\int_0^{\Lambda-\tau} \\ &\quad + \int_{\Theta+\tau}^{2\pi}\int_{\Lambda+\tau}^{2\pi} |\{\cdot\cdot\cdot\}|\, d\omega = I_1 + I_2 + I_3 + I_4 + I_5 + I_6, \end{aligned}$$

where $\delta_{\Theta\Lambda}$ is the function defined in Lemma IV and $\{\cdot\cdot\cdot\}$ replaces the expression

$$\int_0^{\Theta}\int_0^{\Lambda} Q[r_1h(r_2e^{i\theta}, \lambda), r_2e^{i\theta}; \theta', \lambda'] d\theta d\lambda - \delta_{\Theta\Lambda}(\theta', \lambda').$$

By (4. 8),

$$I_1 + I_2 < 2\epsilon;$$

by (3. 6),

$$I_3 + I_4 + I_5 + I_6 < \int_0^{2\pi}\int_0^{2\pi} d\omega \cdot \eta(\tau, 1 - r_1, 1 - r_2).$$

(4. 7) is thus proved.

The following are two immediate consequences of our theorem:

Corollary 1. *If* U *may be represented in the form* (I), ω *is determined by* U *in an unique manner.*

Corollary 2. *If* U *may be represented in the form* (I),

$$(4.9)\qquad \int_0^{2\pi}\int_0^{2\pi} | U[r_1h(r_2e^{i\theta}, \lambda), r_2e^{i\theta}] |\, d\theta d\lambda < c < \infty$$

In fact,

$$\lim_{r_1\to 1;\, r_2\to 1} \int_0^{2\pi}\int_0^{2\pi} \left| U[r_1 h(r_2 e^{i\theta}, \lambda), r_2 e^{i\theta}] \right| d\theta d\lambda = \int_0^{2\pi}\int_0^{2\pi} | d\omega | .$$

(4.9) is the *necessary and sufficient condition* for the possibility of representing a given biharmonic function U in the form (I).

If

$$\int\!\!\int_e | d\omega_{r_1 r_2} | \leq v(\text{mes}\, e), \quad \lim_{\epsilon\to 0} v(\epsilon) = 0,$$

ω is a completely continuous function, i. e. a Lebesgue Integral of a point function u. In this case the representation (I) may be replaced by the following:

$$\text{(II)} \qquad U(z_1, z_2) = \int_0^{2\pi}\int_0^{2\pi} Q(z_1, z_2; \theta, \lambda)\, u(\theta, \lambda)\, d\theta d\lambda.$$

If U is bounded, u is bounded, too.

5. Boundary values of functions which may be represented in the form (I).

In this section we consider an arbitrary function U which may be represented in the form (I), i. e. we do not suppose U to be biharmonic. We shall prove the following

THEOREM. *Suppose that* $U(z_1, z_2)$ $(\{z_1, z_2\} \in \mathfrak{M})$ *may be represented in the form* (I). *If the c. a. set function* $\omega(e)$ *possesses in the point* $\{\Theta, \Lambda\}$ *a finite strong derivative* D, *i. e. if*

$$\lim_{\sigma\to 0,\, \sigma'\to 0,\, \tau\to 0,\, \tau'\to 0} \frac{1}{(\sigma+\sigma')(\tau+\tau')} \int_{\Theta-\sigma'}^{\Theta+\sigma}\int_{\Lambda-\tau'}^{\Lambda+\tau} d\omega = D,$$

U *possesses in the point* $\{h(e^{i\theta}, \Lambda), e^{i\theta}\} \in \mathfrak{F}$ *the sectorial limit* D.

COROLLARY. *If*

$$(5.1) \qquad \begin{cases} \omega(e) = \displaystyle\int\!\!\int_e u(\theta, \lambda)\, d\theta d\lambda, \\ \displaystyle\int_0^{2\pi}\int_0^{2\pi} | u(\theta, \lambda) |^p \, d\theta d\lambda < \infty \qquad (p > 1), \end{cases}$$

U *possesses finite sectorial limits almost everywhere on the distinguished boundary surface* $\mathfrak{F}$.[14]

It is convenient to state first the following lemmas.

---

[14] If the conditions (5.1) are satisfied, ω possesses almost everywhere a finite strong derivative, as was proved by Zygmund, see Zygmund (1). A theorem analogous to our corollary was proved by Bergman and Marcinkiewicz under the hypothesis that U is biharmonic. Their method is entirely different from the one used here.

LEMMA V. *Suppose that* $g(\theta)$ *is a function of bounded variation,*

$$\left|\frac{g(\theta)}{\theta}\right| < c \text{ for } |\theta| \leq \tau < \frac{\pi}{2}, \int_0^{2\pi} |dg(\theta)| \leq m \tag{5.2}$$

and set

$$u(z) = \int_{-\pi}^{\pi} P(z, \theta) dg(\theta).$$

There exist for $a > 1$ *two numbers*

$$k(a) \text{ and } \delta(m, \tau, a, c)$$

such that

$$|u(z)| < ck(a) \text{ for } |z - 1| \leq \delta(m, \tau, a\, c)$$

and

$$|1 - z| < a(1 - |z|). \tag{5.3}$$

LEMMA VI. *Suppose that* $g(\lambda)$ *is a function of bounded variation, satisfying* (5.2) *and*

$$u(z_1; z_2) = \int_{-\pi}^{\pi} \bar{P}(z_2; z_1, \lambda) dg(\lambda). \tag{5.4}$$

There exist for $a > 1$ *two numbers*

$$\bar{k}(a) \text{ and } \bar{\delta}(m, \tau, a, c)$$

such that

$$|u(z_1, z_2)| < c\bar{k}(a)$$

for

$$|z_1 - h(z_2, 0)| < \bar{\delta}(m, \tau, a, c)$$

and

$$|z_1 - h(z_2, 0)| < a \min_{0 \leq \Lambda \leq 2\pi} |z_1 - h(z_2, \Lambda)|. \tag{5.5}$$

Proof of Lemma V. We have

$$u(z) = \int_{-\pi}^{-\tau} Pdg + \int_{-\tau}^{\tau} Pdg + \int_{\tau}^{\pi} Pdg = I_1 + I_2 + I_3,$$

For sufficiently small $|1 - z|$,

$$|I_1 + I_3| < \frac{1}{2\pi} \frac{1 - |z|}{|z|(1 - \cos \tau/2)} m. \tag{5.6}$$

On the other hand,

$$I_2 = [g(\tau) - g(-\tau)]P(z, \tau) - \int_{-\tau}^{\tau} g(\theta) \frac{\partial P(z, \theta)}{\partial \theta} d\theta.$$

Write

$$z = re^{i\phi}, \quad I(z, \tau) = \int_{-\tau}^{\tau} \left|\theta \frac{\partial P(z, \theta)}{\partial \theta}\right| d\theta.$$

For sufficiently small $1-r$

$$|I_2| < m\,\frac{1}{2\pi}\,\frac{1-r}{r(1-\cos\tau/2)} + cI(z,\tau). \tag{5.7}$$

If z satisfies the condition (5.3) and $1-r$ is sufficiently small, we have

$$|\phi| < \tau \text{ and } \phi/(1-r) < k',$$

where k' is a constant which depends only upon a. Since

$$\frac{\partial P(z,\theta)}{\partial\theta} = \frac{1}{2\pi}\,\frac{2(1-r^2)\,r\sin(\theta-\phi)}{[1-2r\cos(\theta-\phi)+r^2]^2},$$

$$\theta\frac{\partial P}{\partial\theta}\begin{cases}\leq 0 \text{ for } -\tau\leq\theta\leq 0\leq\phi\\ \geq 0 \text{ for } 0\leq\theta\leq\phi\\ \leq 0 \text{ for } 0\leq\phi\leq\theta\leq\tau.\end{cases}$$

We may suppose that $\phi\geq 0$ so that

$$\begin{aligned}I(z,\tau) &= -\int_{-\tau}^{0}\theta\frac{\partial P}{\partial\theta}\,d\theta + \int_{0}^{\phi}\theta\frac{\partial P}{\partial\theta}\,d\theta - \int_{\phi}^{\tau}\theta\frac{\partial P}{\partial\theta}\,d\theta\\ &= -[\theta P]_{-\tau}^{0} + \int_{\tau}^{0} Pd\theta + [\theta P]_{0}^{\phi} - \int_{0}^{\phi} Pd\theta\\ &\quad -[\phi P]_{\phi}^{\tau} + \int_{\phi}^{\tau} Pd\theta < 1 + 2\phi P(z,\phi)\\ &= 1 + \frac{1+r}{\pi}\cdot\frac{\phi}{1-r} < 1 + \frac{2k'}{\pi}.\end{aligned}$$

From this inequality our theorem follows (by (5.6) and (5.7)).

Proof of Lemma VI. From (5.5) follows:

$$\begin{aligned}&|W(z_1;z_2)-1| < M\,|z_1-h(z_2,0)|\\ &\qquad < aM\min_{0\leq\Lambda\leq 2\pi}|z_1-h(z_2,\Lambda)| < aM^2[1-|W(z_1;z_2)|].\end{aligned}$$

We may, therefore, apply the identity (3.1) and Lemma V. We get

$$\bar{k}(a) = KMk(aM^2),\ \bar{\delta}(m,\tau,a,c) = \delta(KMm,\tau/KM,aM^2,KMc).$$

We are now in position to prove our theorem. We may suppose that $\Theta=\Lambda=0$, $D=0$. Write

$$\Psi(\Theta,\Lambda) = \int_{-\pi}^{\Theta}\int_{-\pi}^{\Lambda} d\omega - \int_{-\pi}^{\Theta}\int_{-\pi}^{0} d\omega + \int_{-\pi}^{0}\int_{-\pi}^{0} d\omega - \int_{-\pi}^{0}\int_{-\pi}^{\Lambda} d\omega,$$

then

$$U(z_1,z_2) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} Q(z_1,z_2;\theta,\lambda)\,d_\theta d_\lambda\Psi(\theta,\lambda). \tag{5.8}$$

Set

$$\Psi(\theta,\lambda)=\theta\lambda\psi(\theta,\lambda),\ \int_0^{2\pi}\int_0^{2\pi}d\omega=m.$$

If

$$\{z_1,z_2\}\to\{h(1,0),1\}$$

is a sectorial approach (see § 2), we may choose a positive τ such that

$$|\psi(\theta,\lambda)|<\frac{\epsilon}{k(a)\bar{k}(a)}\text{ for }|\theta|<\tau,|\lambda|<\tau,$$

(5. 8) may be written in the form:

$$U(z_1,z_2)=\int_{-\pi}^{\pi}\bar{P}(z_2,\theta)d_\theta\theta\int_{-\pi}^{\pi}P(z_2;z_1,\lambda)d_\lambda\lambda\psi(\theta,\lambda).$$

If $|z_1-h(1,0)|$ and $|z_2-1|$ are sufficiently small,

$$|z_1-h(z_2,0)|<\bar{\delta}\left[m,\tau,a,\frac{\epsilon}{k(a)\bar{k}(a)}\right]$$

and

$$|z_2-1|<\delta\left[m,\tau,a,\frac{\epsilon}{k(a)}\right],$$

hence (by Lemma VI)

$$\left|\int_{-\pi}^{\pi}\bar{P}(z_2;z_1,\lambda)d_\lambda\lambda\psi(\theta,\lambda)\right|<\frac{\epsilon}{k(a)}$$

and (by Lemma V)

$$|U(z_1,z_2)|<\epsilon.$$

Our theorem is thus proved.

6. Bounded analytic functions. We now apply our results to analytic functions of two complex variables. The real and imaginary parts of an analytic function $f(z_1,z_2)$ are biharmonic functions. Therefore the following theorem is true

THEOREM 1. *Suppose that* $f(z_1,z_2)$ *is a bounded analytic function* $(\{z_1,z_2\}\in\mathfrak{M})$. f *may be represented in the form*

$$f(z_1,z_2)=\int_0^{2\pi}\int_0^{2\pi}Q(z_1,z_2;\theta,\lambda)F(\theta,\lambda)d\theta d\lambda, \tag{6. 1}$$

where F *is a bounded measurable (complex) function.* f *posseses the sectorial limit* $F(\theta,\lambda)$ *at almost every point* $\{h(e^{i\theta},\lambda),e^{i\theta}\}\in\mathfrak{F}$.

In fact, this theorem is an immediate consequence of the theorems of § 4 and § 5.

We are now in position to prove

Theorem 2. *Suppose that $f(z_1, z_2)$ is a bounded analytic function in $\mathfrak{M}$. If for almost every point $\{h(e^{i\theta}, \lambda), e^{i\theta}\}$ of the distinguished boundary surface $\mathfrak{F}$ there exists a sequence $\{z_1^\nu, z_2^\nu\}$ such that*

$$\{z_1^\nu, z_2^\nu\} \epsilon \mathfrak{M}, \{z_1^\nu, z_2^\nu\} \rightarrow \{h(e^{i\theta}, \lambda), e^{i\theta}\} \qquad (\nu \rightarrow \infty),$$
$$| z_1^\nu - h(z_2^\nu, \lambda) | < a \cdot \min_{0 \leq \theta \leq 2\pi} | z_1^\nu - h(z_2^\nu, \Lambda) | ,$$
$$| z_2^\nu - e^{i\theta} | < a(1 - | z_2^\nu |)$$ [15]

and

$$\overline{\lim_{\nu \rightarrow \infty}} | f(z_1^\nu, z_2^\nu) | \leq \phi(\theta, \lambda)$$

where ϕ is a measurable function, then

$$\log | f(z_1, z_2) | \leq \int_0^{2\pi} \int_0^{2\pi} Q(z_1, z_2; \theta, \lambda) \log \phi(\theta, \lambda) d\theta d\lambda \tag{6.2}$$

Proof. By Theorem 1 we may suppose $\phi = | F |$. Suppose that $\{r_1^\nu, r_2^\nu\}$ is a sequence satisfying the conditions (4.1), $\alpha > 0$ and

$$\Psi_\nu^{(a)}(\theta, \lambda) = \log \max \{\alpha, | f[r_1^\nu h(r_2^\nu e^{i\theta}, \lambda), r_2^\nu e^{i\theta}] | \},$$
$$g_\nu^{(a)}(z_1, z_2) = \int_0^{2\pi} \int_0^{2\pi} Q(r_1^\nu, r_2^\nu; z_1, z_2; \theta, \lambda) \Psi_\nu^{(a)}(\theta, \lambda) d\theta d\lambda.$$

By Theorem B (§ 1),

$$\log | f[r_1^\nu h(z_2, \lambda), z_2] | - g_\nu^{(a)}[r_1^\nu h(z_2, \lambda), z_2]$$

is a regular harmonic function for $| z_2 | < r_2^\nu$, with the possible exception of certain points where this expression is negative infinite. For $z_2 = | r_2 |$ this function is non-positive. Hence,

$$\log | f[r_1^\nu h(z_2, \lambda), z_2] | - g_\nu^{(a)}[r_1^\nu h(z_2, \lambda), z_2] \leq 0.$$

In the same way we get

$$\log | f(z_1, z_2) | < g_\nu^{(a)}(z_1, z_2) \qquad (\nu = 1, 2, \cdots).$$

On the other hand,

$$\lim_{\nu \rightarrow \infty} | f[r_1^\nu h(r_2^\nu e^{i\theta}, \lambda), r_2^\nu e^{i\theta}] | = \phi(\theta, \lambda)$$

for almost every point $\{\theta, \lambda\}$, for the approach

$$\{r_1^\nu h(r_2^\nu e^{i\theta}, \lambda), r_2^\nu e^{i\theta}\} \rightarrow \{h(e^{i\theta}, \lambda), e^{i\theta}\}$$

is a sectorial one (see § 2). Therefore,

$$\lim_{\nu \rightarrow \infty} \Psi_\nu^{(a)}(\theta, \lambda) = \log \max\{\alpha, \phi(\theta, \lambda)\}$$

---

[15] a may depend on $\{\theta, \lambda\}$.

(for almost every point $\{\theta, \lambda\}$) and

$$\log |f| \leq \lim_{\nu \to \infty} g_\nu^{(\alpha)} = \int_0^{2\pi} \int_0^{2\pi} Q \log \max \{\alpha, \phi\} d\theta d\lambda.$$

By $\alpha \to 0$ we obtain the desired inequality.

THEOREM 3. *If a bounded analytic function $f(z_1, z_2)$ $(\{z_1, z_2\} \epsilon \mathfrak{M})$ possesses the sectorial limit 0 on a set $E \subset \mathfrak{F}$ of positive exterior measure, $f \equiv 0$.*

Proof. f may be represented in the form (6.1) and possesses almost everywhere on $\mathfrak{F}$ the sectorial limit F. F is measurable. The set $E[F = 0]$ is therefore measurable. By our hypothesis its measure is positive. If we apply the formula (6.2) with $\phi = |F|$, we get $\log |f| = -\infty$, $f \equiv 0$.

If we combine Theorems 1 and 3, we get

THEOREM 4. *Suppose that $f(z_1, z_2)$ $(\{z_1, z_2\} \epsilon \mathfrak{M})$ may be represented in the form $f = g/h$, g and h being bounded analytic functions. f possesses almost everywhere on $\mathfrak{F}$ finite sectorial limits. If these limits are constant on a set E of positive exterior measure, f is constant.*

NEW YORK CITY.

BIBLIOGRAPHY.

S. Bergman.

1. " Über ausgezeichnete Randflächen in der Theorie der analytischen Funktionen von zwei komplexen Veränderlichen," *Mathematische Annalen*, vol. 104 (1931), pp. 611-636.

2. " Über den Wertevorrat einer Funktion von zwei komplexen Veränderlichen," *Mathematische Zeitschrift*, vol. 36 (1932), pp. 171-183.

3. " Zwei Sätze aus dem Ideenkreis des Schwarzschen Lemmas bei den Funktionen von zwei komplexen Veränderlichen," *Mathematische Annalen*, vol. 109 (1934), pp. 324-348.

4. " Über eine in gewissen Bereichen mit Maximumfläche gültige Integraldarstellung von Funktionen zweier komplexen Veränderlichen," *Mathematische Zeitschrift*, vol. 39 (1934), pp. 76-94.

5. " Über eine Integraldarstellung der Funktionen zweier komplexer Veränderlicher," *Recueil Mathématique*, vol. 1 (43) (1936), pp. 851-862.

6. " Sur les fonctions entières et méromorphes de deux variables complexes, I," *Compositio Mathematica*, vol. 3 (1936), pp. 136-173.

7. " Zur Theorie der meromorphen Funktionen von zwei komplexen Veränderlichen," *Recueil Mathématique*, vol. 2 (44), (1937), pp. 609-616.

8. "Über meromorphe Funktionen von zwei komplexen Veränderlichen," *Compositio Mathematica*, vol. 6 (1939), pp. 305-335.

9. "Über die Veranschaulichung der Kreiskörper und der Bereiche mit ausgezeichneter Randfläche," *Jahresbericht DMV*, vol. 42 (1933), pp. 228-252.

S. Bergman and J. Marcinkiewicz.

1. "Sur les fonctions analytiques de deux variables complexes, *Fundamenta Mathematicae*, vol. 33 (1939), pp. 75-94.

L. Bers.

1. "Sur une représentation intégrale des fonctions biharmoniques dans les domaines possédant une surface frontière remarquable," *Comptes rendus*, vol. 208 (1939), pp. 1273-1275.

2. "Sur les valeurs limites des fonctions analytiques de deux variables complexes dans les domaines possédant une surface frontière remarquable," *ibid.*, vol. 208 (1939), pp. 1475-1477.

G. C. Evans.

1. "Sur l'intégrale de Poisson," *ibid.*, vol. 177 (1923), pp. 241-242.

G. C. Evans and H. E. Bray.

1. "Sur l'intégrale de Poisson généralisée," *ibid.*, vol. 176 (1923), pp. 1042-1044.

2. "La formule de Poisson et le problème de Dirichlet," *ibid.*, vol. 176 (1923), pp. 1368-1371.

P. Fatou.

1. "Séries trigonométriques et séries de Taylor," *Acta Mathematica*, vol. 30 (1906), pp. 335-400.

G. Herglotz.

1. "Über Potenzreihen mit positivem Realteil im Einheitskreise," *Leipziger Berichte*, vol. 63 (1911), pp. 501-511.

J. Radon.

1. "Theorie und Anwendung der absolut additiven Mengenfunktionen," *Wiener Berichte*, vol. 122 (1913), pp. 1295-1438.

2. "Über lineare Funktionaltransformationen und Funktionalgleichungen," *ibid.*, vol. 128 (1919), pp. 1083-1121.

F. and M. Riesz.

1. "Über die Randwerte analytischer Funktionen," *4. Congr. scand. Math.*, Stockholm, 1916.

S. Warschawski.

1. "Über das Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung," *Mathematische Zeitschrift*, vol. 35 (1932), pp. 321-456.

A. Zygmund.

1. "On the differentiability of multiple integrals," *Fundamenta Mathematicae*, vol. 23 (1934), pp. 143-149.

ON SETS OF MULTIPLICITY FOR TRIGONOMETRICAL SERIES.*

By R. Salem.

---

1. We shall consider, throughout this paper, symmetrical perfect sets constructed in the closed interval $(0, 2\pi)$ as follows: We divide the interval in three parts of lengths proportional to ξ_1, $1 - 2\xi_1$, ξ_1 respectively. We remove the central open interval and divide each of the two remaining intervals in three parts of lengths proportional to ξ_2, $1 - 2\xi_2$, ξ_2; we remove the two central open intervals, and so on indefinitely, the sequence $\{\xi_k\}$ being such that $0 < \xi \leq \frac{1}{2}$. After p operations, $2^p - 1$ intervals have been removed, which we denote by δ_{pk} $(k = 1, 2, \cdots 2^p - 1)$ and 2^p intervals are left, which we denote by η_{pk} $(k = 1, 2, \cdots 2^p)$ and which are of common length $2\pi\xi_1\xi_2 \cdots \xi_p$. When $p \to \infty$ we obtain a symmetrical perfect set P whose measure is $2\pi.\lim 2^p\xi_1\xi_2 \cdots \xi_p$. Here we shall consider only sets of measure zero.

With every set P we shall associate a monotonic continuous function $F(x)$, constant in every interval contiguous to P, but increasing from one interval to another, by the following well known process. For every p let $F_p(x)$ be a non-decreasing continuous function defined by the following conditions: $F_p(0) = 0$, $F_p(2\pi) = 1$, $F_p(x) = k/2^p$ in δ_{pk} $(k = 1, 2, \cdots 2^p - 1)$, and $F_p(x)$ linear in each η_{pk} $(k = 1, 2, \cdots 2^p)$. $F(x)$ will be the limit of $F_p(x)$ as p tends to infinity.

Finally we shall denote by c_n the Stieltjes integral $\int_0^{2\pi} e^{nix}dF$ (n integral) which is the Fourier-Stieltjes coefficient of $2\pi dF$. It is not difficult to show[1] that

$$c_n = e^{\pi n i} \prod_{k=1}^{\infty} \cos \pi n\, \xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k)$$

It is well known that if $c_n = o(1)$ the set P is a set of multiplicity for trigonometrical series.

---

* Received July 21, 1941.

[1] This result has been proved by Carleman, "Sur les équations intégrales singulières à noyau réel et symétrique," Uppsala (1923); see also Hille & Tamarkin, *American Mathematical Monthly*, vol. 36 (1929), pp. 255-264; and R. Salem, "On absolute convergence of trigonometrical series," *Duke Mathematical Journal*, vol. 8 (1941), pp. 317-334.

2. Let us consider now the set of all sets P for which $a_k \leq \xi_k \leq b_k$ $\{a_k\}$ and $\{b_k\}$ being two fixed sequences such that $0 < a_k < b_k < \frac{1}{2}$. If we put

$$(1) \qquad \xi_k = (b_k - a_k)\eta_k + a_k$$

and if $a_k \leq \xi_k \leq b_k$ then $0 \leq \eta_k \leq 1$. Following Steinhaus[2] we can map the interval $0 \leq t \leq 1$ on the "cube" $0 \leq \eta_k \leq 1$ $(k = 1, 2, \cdots)$ of a space of infinitely many dimensions, in the following way: if

$$t = \cdot\, \alpha_1\alpha_2\alpha_3 \cdots$$

is the infinite dyadic development of t we put

$$\begin{aligned} \eta_1(t) &= \cdot\, \alpha_1\alpha_3\alpha_6\alpha_{10} \cdots \\ \eta_2(t) &= \cdot\, \alpha_2\alpha_5\alpha_9\alpha_{14} \cdots \\ \eta_3(t) &= \cdot\, \alpha_4\alpha_8\alpha_{13}\alpha_{19} \cdots \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned}$$

With the exception of the points t for which one at least of these developments is finite (such points form a set of measure zero) the correspondence thus obtained is a one-one correspondence. Moreover, it has been shown by Steinhaus that if $\Phi(\eta_1\eta_2 \cdots \eta_p)$ is any measurable function of $\eta_1 \cdots \eta_p$, we have

$$(2) \quad \int_0^1 \Phi[\eta_1(t), \eta_2(t) \cdots \eta_p(t)]\,dt = \int_0^1 \int_0^1 \cdots \int_0^1 \Phi(\eta_1\eta_2 \cdots \eta_p)\,d\eta_1 \cdots d\eta$$

whenever either side exists.

We have thus established a correspondence between the sets P for which $a_k \leq \xi_k \leq b_k$ and the values of the single variable t in the interval $(0, 1)$. If a certain property belongs to all P, with the possible exception of the P corresponding to a set of t which is of measure zero, we shall say that this property belongs to almost all P.

The purpose of this paper is to prove the following theorems:

Let us consider all the symmetrical perfect sets P *for which* $a_k \leq \xi_k \leq b_k$, *where* $0 < a_k < b \leq \frac{1}{2}$ *and where the sequences* $\{a_k\}$, $\{b_k\}$ *are not "too near" each other, namely, where* $b_p - a_p \geq 1/\omega(p)$, $\omega(p)$ *being an increasing sequence such that* $\log \omega(p) = o(p)$. *Then*:

THEOREM I. *If* $\varliminf (a_1a_2 \cdots a_p)^{1/p} = \alpha > 0$, *almost all sets* P *are sets of multiplicity.*

---

[2] *Math. Zeitschrift*, vol. 21 (1929), pp. 408-416 and *Studia Mathematica*, vol. 2 (1930), pp. 21-39. See also Kaczmarz and Steinhaus, *Theorie der Orthogonalreihen*, Warszowa-Lwów, 1935, p. 134.

THEOREM II. *Moreover, on the same assumption, there exists a constant* s_0 *depending only on* α, *and such that, for every* $s > s_0$ *the series* $\Sigma \mid c_n \mid^s$ *converges for almost all sets* P.

THEOREM III. *Furthermore, on the same assumption, there exists a constant* δ_0 *depending only on* α, *and such that, for every* $\delta < \delta_0$, *the relation* $c_n = o(1/n^\delta)$ *holds good for almost all sets* P.

THEOREM IV. *Theorem* I *remains true if instead of supposing* $\underline{\lim}\, (a_1 a_2 \cdots a_p)^{1/p} = \alpha > 0$, *we suppose only that* $\underline{\lim}\ p^{\frac{1}{2}-\beta}(a_1 a_2 \cdots a_p)^{1/p} > 0$ β *being any positive number, as small as we please.*

It should be observed that we can always choose the sequence $\{b_k\}$ such that *all* the sets considered are sets of measure zero (provided, of course, that the set for which $\xi_k = a_k$ is itself of measure zero).

Theorem III is related to a problem which has been solved first by Littlewood,[3] namely the problem of finding a monotonic singular function with Fourier-Stieltjes coefficients of order smaller than $1/n^\sigma$, $\sigma > 0$. Very complete results have been obtained further in this direction by Wiener and Wintner[4] for monotonic singular functions, which do not seem, however, to be of the Cantor type considered here.[5]

Theorem IV is a generalization of Theorem I which is interesting on account of the following fact. It can be proved[6] that any symmetrical perfect set for which

$$(\xi_1 \xi_2 \cdots \xi_p)^{1/p} = o(1/\sqrt{p})$$

is such that $\overline{\lim} \mid c_n \mid = 1$.

3. The proof of the above theorems is based on the following lemma:

l being a positive number ≥ 1 *and s any positive number, we have:*

---

[3] *Quarterly Journal of Mathematics*, vol. 7 (1936), pp. 219-226.

[4] *American Journal of Mathematics*, vol. 60 (1938), pp. 513-532, and *Journal of Mathematics and Physics*, vol. 17 (1938), pp. 233-246. See also Schaeffer, *American Journal of Mathematics*, vol. 61 (1939), p. 934.

[5] A singular monotone function is said to be of the Cantor type if it is constant in every interval contiguous to a perfect set of measure zero. For the order of the Fourier-Stieltjes coefficients of such functions see a paper of the author which is about to appear in the *Journal of Mathematics and Physics.*

[6] See the above-quoted paper "On absolute convergence of trigonometrical series," p. 326.

$$\int_0^1 |\cos(l\pi x + m)|^s\, dx \le (1 + 1/[l])\, 1/\pi \int_0^\pi |\cos y|^s\, dy$$

where $[l]$ *is the greatest integer contained in* l.

The proof is immediate: by putting $l\pi x + m = y$ the first integral becomes

$$\frac{1}{l\pi}\int_m^{m+l\pi} |\cos y|^s\, dy \le \frac{1}{l\pi}\int_m^{m+[l]\pi+\pi} |\cos y|^s\, dy$$
$$= \frac{[l]+1}{l\pi}\int_0^\pi |\cos y|^s\, dy \le \left(1 + \frac{1}{[l]}\right)\frac{1}{\pi}\int_0^\pi |\cos y|^s\, dy.$$

We shall put $\epsilon_s = \frac{1}{\pi}\int_0^\pi |\cos y|^s\, dy$.

ϵ_s decreases from 1 to zero, as s increases from 0 to ∞.

We can now prove our theorems. Let s be a positive number; we have

$$|c_n|^s = \prod_{k=1}^{\infty} |\cos \pi n \xi_1 \xi_2 \cdots \xi_{k-1}(1 - \xi_k)|^s$$
$$\le \prod_{k=1}^{p} |\cos \pi n \xi_1 \cdots \xi_{k-1}(1 - \xi_k)|^s = f_{n,p}.$$

The ξ_k being functions of the η_k by (1) and the η_k functions of t as explained above, we have, by (2)

$$\int_0^1 f_{n,p}\, dt = \int_0^1 \cdots \int_0^1 f_{n,p}\, d\eta_1\, d\eta_2 \cdots d\eta_p$$
$$= \int_0^1 \cdots \int_0^1 f_{n,p-1}\, d\eta_1 \cdots d\eta_{p-1} \int_0^1 |\cos \pi n \xi_1 \cdots \xi_{p-1}(1 - \xi_p)|^s\, d\eta_p$$

The last integral with respect to the variable η_p is equal to

$$J_p = \int_0^1 |\cos \pi n \xi_1 \cdots \xi_{p-1}[1 - a_p - (b_p - a_p)\eta_p]|^s\, d\eta_p.$$

Let us remember now that $\xi_k \ge a_k$ and $b_p - a_p \ge 1/(\omega(p))$ and let us suppose that p is so chosen that

$$(3) \qquad n a_1 a_2 \cdots a_{p-1}(1/\omega(p)) \ge p^2$$

then, by the lemma

$$J_p \le (1 + 1/p^2)\epsilon_s$$

and

$$\int_0^1 f_{n,p}\, dt \le (1 + 1/p^2)\epsilon_s \int_0^1 f_{n,p-1}\, dt$$

But if (3) holds good for $p = q$, it holds good *a fortiori* for $p < q$; hence applying the same argument we have

$$\int_0^1 f_{n,p} dt \leq A\epsilon_s^p$$

A being an absolute constant, and, *a fortiori*

$$\int_0^1 |c_n|^s \, dt \leq A\epsilon_s^p \tag{4}$$

p being any integer such that (3) be satisfied.

In the hypothesis of Theorem I we have, α' being any fixed positive number such that $\alpha' < \alpha$, and p being large enough

$$a_1 a_2 \cdots a_{p-1} \geq \alpha'^{p-1}.$$

and thus, in order to satisfy the inequality (3), it is sufficient to take p such that

$$\log n + (p - 1)\log \alpha' - \log \omega(p) \geq 2\log p$$

and, as $\log \omega(p)/p \rightarrow 0$ it is sufficient to take, when n is large enough

$$p = p(n) = [\theta'(\log n/\log 1/\alpha')] = [\theta(\log n/\log 1/\alpha)]$$

θ being, like θ', any fixed number smaller than 1, as we can take α' as near to α as we please.

θ being thus fixed we have, by (4)

$$\int_0^1 |c_n|^s dt \leq A\epsilon_s^{p(n)} \leq A\epsilon_s^{\theta(\log n/\log 1/\alpha)-1}$$

Now let us take s such that

$$\epsilon_s^{\theta(\log n/\log 1/\alpha)} \leq 1/n^{1+\rho} \qquad (\rho > 0)$$

that is to say

$$\theta\,(\log n/\log 1/\alpha)\log \epsilon_s \leq -(1 + \rho)\log n$$

or

$$\epsilon_s \leq \alpha^{(1+\rho)/\theta}$$

which is possible, ϵ_s tending to zero for $s = \infty$. s being any number thus fixed, the series $\Sigma \int_0^1 |c_n|^s \, dt$ converges so that the series $\Sigma\, |c_n|^s$ converges for almost every t.

It follows, in particular, that $c_n = o(1)$ for almost all sets P, and that consequently almost all sets P are sets of multiplicity, which proves Theorem I.

Let us now define s_0 by the equation

$$\epsilon_{s_0} = 1/\pi \int_0^\pi |\cos y|^{s_0} \, dy = \alpha \tag{5}$$

Then, as $1 + \rho$ and θ can be taken as near to 1 as we please, we get the following result:

For almost all sets P, in the hypothesis of Theorem I, *the series* $\Sigma |c_n|^s$ *converges for* $s > s_0$, s_0 *being given by* (5). This proves Theorem II.

In order to prove Theorem III, let us choose s such that

$$\epsilon_s^{\theta(\log n/\log 1/\alpha)} \leq 1/n^{1+\rho+\gamma} \qquad (\rho > 0, \gamma > 0)$$

which is equivalent to $\epsilon_s \leq \alpha^{(1+\rho+\gamma)/\theta}$.

s being thus chosen, the series $\Sigma n^\gamma \int_0^1 |c_n|^s \, dt$ converges, hence the series $\Sigma n^\gamma |c_n|^s$ converges for almost all P.

As before, let us define s_γ by the equation

$$\epsilon_{s_\gamma} = 1/\pi \int_0^\pi |\cos y|^{s_\gamma} \, dy = \alpha^{1+\gamma} \tag{6}$$

As $1 + \rho$ and θ can be taken as near to 1 as we please, we get the following result:

For almost all sets P, in the hypothesis of Theorem I, *the series* $\Sigma n^\gamma |c_n|^s$ *converges for* $s > s_\gamma$, s_γ *being given by* (6).

Hence we have, for almost all P,

$$c_n = o(1/n^{\gamma/s})$$

for every $s > s_\gamma$.

Now, when γ, which is arbitrary, increases from 0 to ∞, s_γ increases from s_0 to ∞; but by a classical theorem $\epsilon_s = O(1/\sqrt{s})$ for $s = \infty$, hence γ/s_γ tends to zero when $\gamma \to \infty$.

Hence γ/s_γ has actually a maximum, which is the maximum of the expression

$$(1/s \log \alpha) \log 1/\alpha\pi \int_0^\pi |\cos y|^s \, dy$$

when $s_0 < s < \infty$. Let $\delta_0(\alpha)$ be this maximum; then for almost all sets P we have

$$c_n = o(n^{-\delta})$$

for $\delta < \delta_0(\alpha)$, which proves Theorem III.

Remark. It is seen immediately that if α increases, $\delta_0(\alpha)$ increases. Let us suppose that the $\{a_k\}$ are such that

$$\lim\ (a_1 a_2 \cdots a_p)^{1/p} = \tfrac{1}{2} \tag{7}$$

Then, by (5), $s_0 = 2$, the series $\Sigma \mid c_n \mid^s$ converges for $s > 2$ for almost all sets P; and

$$\delta_0(\tfrac{1}{2}) = \underset{2<s<\infty}{\mathrm{Max}} \left(- 1/s \log 2\right) \log 2/\pi \int_0^\pi \mid \cos y \mid^s dy$$

and it is easy to see, by taking $s = 6$, that $\delta_0(\tfrac{1}{2})$ is $> 1/9$.

The case (7) occurs, for example, if $\Sigma(1/\omega(p))$ being a divergent series, we take

$$\tfrac{1}{2}(1 - (2/\omega(k))) \leq \xi_k \leq \tfrac{1}{2}(1 - (1/\omega(k)))$$

4. Proof of Theorem IV. In the hypothesis of this theorem we have

$$(1/p) \log (a_1 \cdots a_p) + (\tfrac{1}{2} - \beta') \log p > 0$$

β' being a positive number $< \beta$.

Hence, in order to satisfy the inequality (3) it is sufficient to take p such that

$$\log n - (\tfrac{1}{2} - \beta')\, p \log p - \log \omega(p) \geq 2 \log p.$$

It is sufficient to take, β'' being a positive number $< \beta'$,

$$(\tfrac{1}{2} - \beta'')\, p \log p \leq \log n$$

This inequality will be satisfied if we take

$$p = p(n) = [(2 + \beta_0) \log n/\log \log n] + 1$$

β_0 being a positive number depending only on β''.

p being thus fixed, we have, by (4)

$$\int_0^1 \mid c_n \mid^s dt \leq A \epsilon_s^{(2+\beta_0)(\log n/\log\log n)}.$$

The difference between the argument here and the proof of Theorem I is that here we shall choose s as a function of n, increasing indefinitely with n. Taking into account that $\epsilon_s \sim (2/\pi s)^{\frac{1}{2}}$ for $s = \infty$ we can write

$$\int_0^1 \mid c_n \mid^s dt \leq (1/s)^{(1+\beta_1)(\log n/\log\log n)}$$

β_1 being a positive number.

Now let us choose s such that

$$(1/s)^{(1+\beta_1).(\log n/\log\log n)} \leq 1/n^{1+\beta_1/2}$$

that is to say

$$(1+\beta_1)(\log s/\log\log n) \geqq 1+\beta_1/2.$$

It is sufficient to take

$$s = s_n = s(n) = (\log n)^\theta$$

θ being a fixed number (depending only on β_1) and *smaller than* 1.

s_n being thus chosen, the series

$$\Sigma n^{\beta_1/3} \int_0^1 | c_n |^{s_n} dt$$

will converge, hence the series

$$\Sigma n^{\beta_1/3} | c_n |^{s_n}$$

will converge almost everywhere. Hence, for almost all sets P

$$(| c_n | \cdot n^{\beta_1/3s_n})^{s_n} = o(1)$$

that is to say, as $s_n \to \infty$ with n

(8) $$| c_n | n^{\beta_1/3s_n} < 1.$$

But $(\beta_1/3s_n) \log n = (\beta_1/3)(\log n)^{1-\theta}$ increases indefinitely with n, as $\theta < 1$; hence $n^{\beta_1/3s_n}$ increases indefinitely with n, and the inequality (8) shows that $c_n = o(1)$ for almost all sets P, which proves Theorem IV.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

ELEMENTARY NOTE ON PRIME NUMBER PROBLEMS OF VINOGRADOFF'S TYPE.*

By R. D. James and H. Weyl.

1. Stating the problem. Vinogradoff's asymptotic formula slightly amended. By his ingenious methods Vinogradoff succeeded in establishing an asymptotic formula for the number of decompositions of a large integer N into three primes, $N = p_1 + p_2 + p_3$. When one follows the exposition in his paper, *Recueil Mathématique*, vol. 2 (44) (1937), pp. 179-195, one finds that he introduces another number N_1 (p. 192), and writing k instead of N_1 one sees that he actually determines the solutions of the equation $p_1 + p_2 + p_3 = k$ by primes $p_i \leqq N$. In this form the problem admits of an immediate generalization[1]: Let $r \geqq 3$ non-vanishing integers $a_1, \cdots, a_r$ and r positive real numbers $\beta_1, \cdots, \beta_r$ be given once for all. We ask for the number $A(N, k)$ of representations of an integer k in the form

$$k = a_1 p_1 + \cdots + a_r p_r \tag{1}$$

by means of primes

$$p_1 \leqq \beta_1 N, \cdots, p_r \leqq \beta_r N. \tag{2}$$

The problem thus stated makes sense even if some of the coefficients a_i are negative. The integer k and the real number $N \geqq 2$ are considered as variables. The letter C stands throughout for a positive constant, not always the same, which is independent of N and k; if other variables enter, indices attached to C will indicate on which of these variables C depends. The main result is a formula

$$A(N, k) = R(N, k) \cdot \mathfrak{S}(k) + (N^{r-1}/n^r) \cdot \rho \tag{I}$$

in which the relative error ρ, with $N \to \infty$ and uniformly with respect to k, tends to zero faster than any negative power n^{-g} of $n = \log N$. The factor

---

* Received June 19, 1941.

[1] J. G. van der Corput has published a number of results concerning this problem, *C. R. Ac. Sci. Paris*, vol. 205 (1937), pp. 479-481 and 591-592; *Nederlandsche Akademie van Wetenschappen, Proceedings*, vol. 41 (1938), p. 80. He even imposes upon the primes p_i congruence restrictions modulo a given integer. But as he mentions existence theorems only, a full account of the asymptotic situation might help "to put the house in order."

$R(N,k)$ takes account of the "quantitative" aspect of k (infinite prime spot) in so far as $R(N,x)$ is defined for arbitrary real values of x, whereas the singular series $\mathfrak{S}(k)$ is an arithmetical function of k alone. For any positive integer q we set

$$q_i = q/(q, a_i), \qquad q^* = q/(q,k),$$

and by means of the Möbius and Euler functions $\mu(q)$ and $\phi(q)$ form

$$(3) \quad \chi(q) = (\mu(q_1)/\phi(q_1)) \cdots (\mu(q_r)/\phi(q_r)) \cdot (\mu(q^*)/\phi(q^*)) \cdot \phi(q).$$

Then

$$(4) \qquad \mathfrak{S}(k) = \mathfrak{S}(k; a_1, \cdots, a_r) = \sum_{q=1}^{\infty} \chi(q).$$

We notice that $\mathfrak{S}(k; a_1, \cdots, a_r)$ depends symmetrically on $k; a_1, \cdots, a_r$, although it should not be forgotten that k may take the value 0. The quantitative factor R may be given the following form. Using the abbreviation $e^{2\pi i x} = e(x)$ we introduce the following functions of a real variable z:

$$(5) \qquad J_i(z) = \int_{\sqrt{N}}^{\beta_i N} (e(a_i z x)/\log x)\, dx \qquad (i = 1, \cdots, r).$$

Then $R(N,x)$ is the Fourier transform of their product,

$$(6) \qquad R(N,x) = \int_{-\infty}^{+\infty} J_1(z) \cdots J_r(z) \cdot e(-xz)\, dz.$$

The proof depends on Vinogradoff's sums extending over the primes p,

$$S(\eta) = S(N,\eta) = \sum_{p \leqq N} e(\eta p), \qquad S_i(\eta) = S(\beta_i N, a_i \eta),$$

and starts with the exact formula

$$(7) \qquad A(N,k) = \int_0^1 S_1(\eta) \cdots S_r(\eta) e(-k\eta)\, d\eta.$$

Following Vinogradoff's exposition step by step we call attention to a few points where we deviate from him. [1] The lower limit in (5) or in the typical function

$$J(z) = \int_{\sqrt{N}}^{\beta N} (e(zx)/\log x)\, dx$$

has been set at $\sqrt{N}$ rather than at 2. Looking at the context where this function makes its first appearance, one realizes that the results are not affected by this change. One immediate advantage is that the estimates

$$(8) \qquad |J(z)| \leqq 2\beta N/n, \qquad |J(z)| \leqq 2/(\pi n\,|z|)$$

hold for *all* z, hence

$$|J_i(z)| \leqq C \cdot \delta(z) \tag{9}$$

throughout where

$$\delta(z) = \begin{cases} N/n \text{ for } |z| \leqq N^{-1}, \\ 1/(n|z|) \text{ for } |z| \geqq N^{-1}. \end{cases}$$

A more important advantage will presently become evident. [2] The integration in (6) which Vinogradoff restricts within the bounds $\pm n^{3h}/N$ has been extended to $\pm \infty$. Because of the second inequality (8) the value of R is not changed thereby by more than $CN^{r-1}n^{-3h(r-1)}$. It is exactly this step which makes the main term in the formula (I) independent of Vinogradoff's exponent h and gives a result from which h has vanished altogether. [3] It is not hard to see why in the expression of $\chi(q)$ the numbers $q_1, \cdots, q_r$ occur. A rational point is of the form $\eta = m/q$ where m is an element of the set G_q of all $\phi(q)$ prime residues mod q. In the reduced form the fraction $a_1\eta = a_1m/q$ has the denominator q_1 and hence $e(a_1\eta p)$ depends only on the residue class mod q_1 to which p belongs. The following elementary arithmetical observations are to the point. q' being a divisor of q, $q = q'd$, the number of elements m of G_q which are congruent mod q' to a given element m' of $G_{q'}$ has the same value v for each m'. Clearly $v \leqq d$ and $v = \phi(q)/\phi(q')$. In case a'/q' is the reduced form of a fraction a/q we therefore have $v \leqq (a, q)$, *a fortiori*

$$\phi(q) \leqq |a| \cdot \phi(q') \tag{10}$$

if $a \neq 0$, and the sum

$$\sum_{m \in G_q} e(am/q) = v \cdot \sum_{m' \in G_{q'}} e(a'm'/q') = (\phi(q)/\phi(q')) \cdot \mu(q').$$

The last remark finds application for $a = -k$. Because of (10), $\mathfrak{S}(k)$ is majorized by the convergent sum

$$|a_1 \cdots a_r| \cdot \sum_q (\phi(q))^{-r-1}.$$

[4] One crucial point in the estimate for the contribution to the integral (7) of part II of the circumference $0 \leqq \eta \leqq 1$ is the fact that

$$\int_0^1 |S(\eta)|^2 \, d\eta = \pi(N) \leqq N.$$

An equally good result is obtained in our case where all the factors $S_i(\eta)$ are different, by observing that $\int_0^1 |S_1(\eta)| \cdot |S_2(\eta)| \, d\eta$ does not exceed the square root of

$$\int_0^1 |S_1(\eta)|^2 \, d\eta \cdot \int_0^1 |S_2(\eta)|^2 \, d\eta = \pi(\beta_1 N) \cdot \pi(\beta_2 N) \leqq \beta_1\beta_2 \cdot N^2.$$

By (9) we realize that

$$|R(N,x)| \leqq C \cdot N^{r-1}/n^{r}.$$

Hence we set

$$R(N,x) = (N^{r-1}/n^{r}) \cdot R^{*}(N,x)$$

and then find not only $|R^{*}(N,x)| \leqq C$ but also

$$\text{(II)} \qquad |R^{*}(N,x) - R^{*}(N,x')| \leqq C \cdot (|x - x'|/N),$$

which indicates that $R^{*}(N,x)$ varies very slowly over wide stretches of the variable x if N is large. Of course C now means constant with respect to N and x (or N and x/N). It seems reasonable to measure x against N. Let us therefore write $R^{*}(N,xN) = R_N(x)$. Then (6) and the inequality (9) prove that the first $r-2$ derivatives of $R_N(x)$ with respect to x are continuous and satisfy relations

$$\text{(11)} \qquad |R_N^{(\mu)}(x)| \leqq C \qquad (\mu = 0, 1, \cdots, r-2).$$

We have thus arrived at the following two propositions:

THEOREM I. *A formula* (I) *holds in which the singular series* $\mathfrak{S}(k)$ *is defined by* (4) *and the relative error* $\rho = \rho(N,k)$ *satisfies an inequality*

$$|\rho(N,k)| \leqq C_g \cdot n^{-g}$$

however large the integral exponent $g \geqq 0$.

THEOREM II. $R(N,x)$ *is defined for all real* x, *and if we set*

$$R(N,xN) = (N^{r-1}/n^{r}) \cdot R_N(x),$$

then

$$|R_N(x)| \leqq C, \qquad |R_N(x) - R_N(x')| \leqq C \cdot |x - x'|.$$

2. Analysis of the quantitative factor. In this section we give a simple interpretation of $R(N,x)$ which leads to an asymptotic expansion in terms of powers of n^{-1}. Since k is not representable in the form (1) under the conditions (2) unless k/N lies between the limits

$$A = \sum_{i=1}^{r} \tfrac{1}{2}\beta_i(a_i - |a_i|) \qquad \text{and} \qquad A' = \sum_{i=1}^{r} \tfrac{1}{2}\beta_i(a_i + |a_i|)$$

we must expect $R_N(x)$ to be non-negative for all values of x and to vanish outside the finite interval $A \leqq x \leqq A'$. Our new representation will put these facts in evidence.

We introduce a measure $d\sigma$ for infinitesimal areas $d\sigma$ in any of the parallel planes $E(x)$,

$$a_1x_1 + \cdots + a_rx_r = \text{const.} = x$$

of the $(x_1, \cdots, x_r)$-space, such that the volume of a (straight or oblique) cylinder erected over $d\sigma$ and bounded by the planes $E(x)$ and $E(x+dx)$ is given by $d\sigma \cdot dx$. The expression

$$R(N,x) = \int_{-\infty}^{+\infty} \left\{ \int \cdots \int_{B(N)} \frac{e(z(a_1x_1 + \cdots + a_rx_r))}{\log x_1 \cdots \log x_r} dx_1 \cdots dx_r \right\} e(-xz)\,dz$$

where $B(N)$ denotes the block

$$\sqrt{N} \leqq x_1 \leqq \beta_1 N, \quad \cdots, \quad \sqrt{N} \leqq x_r \leqq \beta_r N$$

suggests forming the integral

$$(12) \qquad F(x) = \int (1/\log x_1) \cdots (1/\log x_r)\,d\sigma$$

over the intersection of the plane $E(x)$ with the block $B(N)$. Then

$$R(N,x) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} e(zx')F(x')\,dx' \right\} e(-xz)\,dz,$$

and by Fourier's theorem,

$$R(N,x) = F(x).$$

It is evident that $F(x)$ vanishes outside the interval $AN \leqq x \leqq A'N$. We shall show in the next section that it has a continuous derivative, and thus justify our application of Fourier's theorem. We formulate our result as follows:[2]

THEOREM III. *Idealize the r primes p_i as statistically independent continuous quantities x_i such that $\phi_i(x)\,dx$,*

$$(13) \qquad \phi_i(x) = 1/\log x \quad \textit{for} \quad \sqrt{N} \leqq x \leqq \beta_i N, \quad 0 \quad \textit{elsewhere},$$

is the probability of x_i assuming values between x and $x+dx$, then $R(N,x)\cdot dx$ is the probability that $a_1x_1 + \cdots + a_rx_r$ lies between x and $x+dx$.

This is a familiar device in the theory of probability: The link betwwen the densities of the probability distributions of r independent variables x_i and a linear combination of them is expressed as simple multiplication in terms of the Fourier transforms of the densities.

---

[2] Even before Vinogradoff made his fundamental discovery enabling him to prove the 3-term Goldbach theorem without such generalized Riemann hypotheses as Hardy-Littlewood had used, H. Rademacher gave this "probability expression" of the quantitative factor occuring in the asymptotic formula for the r-term Goldbach problem. *Math. Zeitschr.*, vol. 25 (1926), p. 628. Van der Corput obtained similar integrals in a number of related problems, mostly of non-linear character. *Nederlandsche Akademie van Wetenschappen, Proceedings*, vol. 41 (1938), pp. 25-26; 97-107; 344-349; *Math. Annalen*, vol. 116 (1938), pp. 1-50.

Using the notation $\xi = -\log x$ we may write our formula in this way:

$$R_N(x) = \int \frac{n}{n-\xi_1}\cdots\frac{n}{n-\xi_r}\,d\sigma, \tag{14}$$

the integral extending over the intersection of $E(x)$ with the block

$$B_N\colon \qquad N^{-1/4} \leqq x_1 \leqq \beta_1, \cdots, N^{-1/4} \leqq x_r \leqq \beta_r.$$

The asymptotic series for $R_N(x)$ is obtained by expanding the individual factors of the integrand as power series in n^{-1}:

$$\frac{n}{n-\xi} = \sum_{i=0}^{\infty} \xi^i/n^i,$$

$$\frac{n}{n-\xi_1}\cdots\frac{n}{n-\xi_r} = \sum_{i=0}^{\infty} (1/n^i)\cdot p_i(\xi_1,\cdots,\xi_r) \tag{15}$$

where

$$p_i(\xi_1,\cdots,\xi_r) = \sum \xi_1^{i_1}\cdots\xi_r^{i_r}$$
$$(i_1 \geqq 0,\cdots,i_r \geqq 0, \quad i_1+\cdots+i_r = i).$$

In the individual terms we extend the integration down to $x_i = 0$ instead of stopping at $N^{-1/4}$; the error to be expected will not be appreciably higher than $O(N^{-1/4})$. We therefore form the integrals

$$\gamma_i = \gamma_i(x) = \int_{E'(x)} p_i(\xi_1,\cdots,\xi_r)\cdot d\sigma \tag{16}$$

extending over the intersection $E'(x)$ of the plane $E(x)$ with the block

$$B\colon \qquad 0 \leqq x_1 \leqq \beta_1, \cdots, \quad 0 \leqq x_r \leqq \beta_r,$$

and then arrive at

THEOREM IV. *$R_N(x)$ has the asymptotic expansion*

$$R_N(x) \sim \gamma_0 + \gamma_1 n^{-1} + \gamma_2 n^{-2} + \cdots \tag{17}$$

in the sense that the absolute error committed by breaking off after the g-th term $\gamma_{g-1}n^{-(g-1)}$ is less than $C_g n^{-g}$, the constant C_g being independent of N and x.

For $A(N,k)$ we find the asymptotic formula[3]

$$A(N,k) \sim (N^{r-1}/n^r)\mathfrak{S}(k)(\gamma_0(\kappa) + \gamma_1(\kappa)n^{-1} + \cdots), \tag{18}$$

---

[3] Vinogradoff has merely the first term of this asymptotic expansion, and the error he gives is $O(\log n/n)$ instead of $O(1/n)$. In the paper cited above, he follows an intricate and indirect approach. Much simpler is the method given on pp. 55-57 of his paper in *Travaux de l'Institut Mathématique de Tbilissi*, vol. 3 (1938). But he uses the formula (6) instead of our (12).

uniformly valid for all k; the letter κ stands for k/N. The thoroughly elementary proof follows in the next section. $R_N(x)$ and $\gamma_i(x)$ vanish outside the interval $A \leqq x \leqq A'$, $R_N(x)$ is non-negative throughout, and the same holds for $\gamma_i(x)$ provided $\beta_1 \leqq 1, \cdots, \beta_r \leqq 1$.

The first coefficient γ_0 is of particular interest. It equals the area of the $(r-1)$-dimensional convex polygon $E'(x)$. Using the method of convolution explained in the next section, we obtain by induction with respect to r this fairly explicit result: Put $\alpha_i = a_i\beta_i$ and for any function $f(x)$ denote the difference $f(x) - f(x - \alpha_i)$ by $\Delta_i f(x)$. Introduce

$$U(x) = \begin{cases} 0 & \text{for} \quad x \leqq 0, \\ \dfrac{x^{r-1}}{(r-1)\,!} & \text{for} \quad x \geqq 0. \end{cases}$$

Then

$$\gamma_0(x) = \frac{1}{a_1 \cdots a_r} \; \Delta_1 \cdots \Delta_r U(x).$$

We add a few obvious observations concerning the question under which circumstances some of the inequalities (2) are redundant.

If all $a_1, \cdots, a_r$ are positive, one will ask for the number of representations of a large integer N in the form

$$N = a_1p_1 + \cdots + a_rp_r$$

by means of r primes p_i, without introducing any upper bounds for the p_i. However, one may add the conditions

$$p_1 \leqq N/a_1, \cdots, p_r \leqq N/a_r$$

without altering the problem. Hence it falls under our scheme. The number $A(N)$ of representations equals $(N^{r-1}/n^r)R_N \cdot \mathfrak{S}(N)$ with a relative error of the type ρ mentioned in Theorem I, and for R_N and the coefficients γ_i of its asymptotic expansion we have expressions which arise from (14) and (16), respectively, by setting $x = 1$ and replacing the blocks B_N and B by the "infinite" blocks

$$\begin{aligned} B_N^{(0)}&: \quad x_1 \geqq N^{-\frac{1}{2}}, \cdots, x_r \geqq N^{-\frac{1}{2}}, \\ B^{(0)}&: \quad x_1 \geqq 0, \cdots, x_r \geqq 0 \end{aligned}$$

respectively. In particular,

$$\gamma_0 = \frac{1}{a_1 \cdots a_r} \cdot \frac{1}{(r-1)\,!}$$

If $a_1, \cdots, a_s$ are positive and $a_{s+1}, \cdots, a_r$ negative, then the equation (1) together with the $r - s$ inequalities

$$(2_s) \qquad p_{s+1} \leqq \beta_{s+1}N, \cdots, p_r \leqq \beta_r N$$

results in the following bounds for $p_1, \cdots, p_s$:

$$p_i \leqq \frac{x_0 - \alpha_{s+1} - \cdots - \alpha_r}{a_i} N \qquad (i = 1, \cdots, s)$$

as long as $k \leqq x_0 N$. Hence the number $A(N, k)$ of solutions of (1) in primes satisfying the inequalities (2_s) is asymptotically given by a formula (18) where $\gamma_i(x)$ is the integral (16) extending over the intersection of $E(x)$ with the partially infinite block

$$x_1 \geqq 0, \cdots, x_s \geqq 0, \qquad 0 \leqq x_{s+1} \leqq \beta_{s+1}, \cdots, \qquad 0 \leqq x_r \leqq \beta_r.$$

Its validity, however, is uniform only for $\kappa \leqq x_0$ if x_0 is any preassigned number. In particular,

$$\gamma_0(x) = \frac{1}{a_1 \cdots a_r} \cdot \Delta_{s+1} \cdots \Delta_r U(x).$$

Examples: 1) $p_1 + p_2 - p_3 = k, \qquad p_3 \leqq N.$

$$\begin{array}{c|c|c} \gamma_0(x) = \frac{1}{2} + x & \frac{1}{2}(1+x)^2 & 0 \\ \hline \text{for } x \geqq 0 & -1 \leqq x \leqq 0 & x \leqq -1 \end{array}.$$

2) $\qquad p_1 + p_2 - p_3 = k; \qquad p_1 \leqq N, \quad p_2 \leqq N.$

Write the equation in the form $p_3 - p_1 - p_2 = -k$.

$$\begin{array}{c|c|c|c} \gamma_0(x) = 1 & 1 - \frac{1}{2}x^2 & \frac{1}{2}(2+x)^2 & 0 \\ \hline \text{for } x \geqq 0 & -1 \leqq x \leqq 0 & -2 \leqq x \leqq -1 & x \leqq -2 \end{array}.$$

3. Proofs of Theorems III and IV. We have to investigate the integral

$$F(x) = \int_{E(x)} \phi_1(x_1) \cdots \phi_r(x_r) \cdot d\sigma,$$

where the ϕ_i are defined by (13). Let us, more generally, discuss the case of arbitrary densities ϕ_i each vanishing outside a finite interval. Using the x_1-axis as the direction in which we erect a cylinder over the area $d\sigma$ on $E(x)$ we obtain

$$F(x) = \frac{1}{|a_1|} \cdot \int_{-\infty}^{+\infty} \cdots \int \phi_1\left(\frac{x - a_2x_2 - \cdots - a_rx_r}{a_1}\right) \cdot \phi_2(x_2) \cdots \phi_r(x_r)\, dx_2 \cdots$$

Write

$$\phi^{[i]}(x) = \frac{1}{|a_i|} \cdot \phi\left(\frac{x}{a_i}\right)$$

and introduce the linear, commutative and distributive operation o of convolution by

$$f_1 \circ f_2 = \int_{-\infty}^{+\infty} f_1(x-y) f_2(y)\,dy = \int_{-\infty}^{+\infty} f_2(x-y) f_1(y)\,dy.$$

Then

$$F = \phi_1^{[1]} \circ \phi_2^{[2]} \circ \cdots \circ \phi_r^{[r]}.$$

The notation $\times \phi_1 \times \phi_2 \times \cdots \times \phi_r$ will also be used for this convolution. But then a permutation of the factors must be accompanied by the same permutation of the numbers $a_1, \cdots, a_r$ on which the definition of the $\times$ multiplication is based.

Let f_1, f_2 denote functions vanishing outside a finite interval.

LEMMA 1. *If f_1, f_2 are piecewise continuous, then $f = f_1 \circ f_2$ is continuous. $|f| \leqq |f_1| \circ |f_2|$, and $|f_1(x)| \leqq M_1$ implies $|f(x)| \leqq M_1 \cdot \int |f_2|\,dx$.*

Suppose f_1 continuous and f_2 of bounded variation, and set

$$g = (f_1, df_2) = \int_{-\infty}^{+\infty} f_1(x-y) \cdot df_2(y).$$

If

$$\int_{-\infty}^{x} f_1 dx = F_1(x), \qquad \int_{-\infty}^{x} g dx = G(x),$$

then

$$G = (F_1, df_2) = f_1 \circ f_2.$$

Hence:

LEMMA 2. *If f_1 is continuous and f_2 of bounded variation, then $g = (f_1, df_2)$ is continuous, and $|f_1(x)| \leqq M_1$ implies $|g(x)| \leqq M_1 \cdot \int |df_2|$. The derivative of $f = f_1 \circ f_2$ is g, and if f_1 has a continuous derivative f'_1 then $g' = (f'_1, df_2)$.*

From these lemmas we deduce that our $F = f_1 \circ \cdots \circ f_r$ whose factors are $f_i = \phi_i^{[i]}$, ϕ_i being defined by (13), has the continuous derivative

$$F' = (f_1 \circ \cdots \circ f_{r-1}), df_r.$$

One may even continue derivation up to the order $r-2$:

$$F^{(r-2)} = (f_1 \circ f_2), df_3, \cdots, df_r.$$

The $(r-1)$-th derivative may still be written in a similar form but it has discontinuities at the points

$$x_1 = a_1\rho_1 + \cdots + a_r\rho_r \qquad (\rho_i = \delta_i \text{ or } \delta'_i) \tag{19}$$

where $\delta_i = \sqrt{N}$, $\delta'_i = \beta_i N$ are the discontinuities of $\phi_i(x)$. One may readily write down the jumps of $F^{(r-1)}$ at the points (19) in terms of the jumps $\Delta_i(\rho_i)$ of the individual functions ϕ_i. As a matter of fact, our $R = F$ is piecewise *analytic*, the pieces being sewed together at the seams (19) in such a fashion that the function itself and its derivatives up to the order $r-2$ go over continuously. We may also confirm (11) by this method, even including $\mu = r-1$.

The proof of Theorem IV would hardly need any explanation were $\xi = -\log x$ bounded in the intervals in which it is used. This being not the case,[4] we feel obliged to go into details. First determine the remainder P_g of the expansion (15) left over when one breaks off after the g-th term. Applying to each factor $n/(n-\xi_i)$ in succession the finite geometric progression broken off after a suitable term, with its well-known remainder, we find that P_g equals the sum of all terms of the following form

$$n^{-g} \cdot \xi_1^{i_1} \cdots \xi_{s-1}^{i_{s-1}} \cdot \xi_s^{h} \frac{n}{n-\xi_s} \cdot \frac{n}{n-\xi_{s+1}} \cdots \frac{n}{n-\xi_r}$$

where $1 \leqq s \leqq r$, $i_1 + \cdots + i_{s-1} = i < g$ and $h = g - i$. An additional feature is the extension of the integration down to the point $x = 0$ instead of $N^{-1/4}$ in the terms of the approximating sum. Therefore the following situation prevails. We denote by $l^*(x)$ the typical function equaling

$$n/(n-\xi) \text{ for } N^{-1/4} \leqq x \leqq \beta, \qquad 0 \text{ elsewhere,}$$

and by $l(x)$ any function

$$\xi^g \text{ for } 0 \leqq x \leqq \beta, \qquad 0 \text{ elsewhere,}$$

with an integral exponent $g \geqq 0$. This particular $l(x)$ is said to be of order g. Moreover we introduce a function $L(x)$ of order $g \geqq 1$ which equals

$$\xi^g n/(n-\xi) \text{ for } N^{-1/4} \leqq x \leqq \beta, \qquad 0 \text{ elsewhere,}$$

and a function $\Lambda(x)$ of order $g \geqq 0$ equaling

$$-\xi^g/n^g \text{ for } 0 \leqq x \leqq N^{-1/4}, \qquad 0 \text{ elsewhere.}$$

Then the difference

$$R_N(x) - (\gamma_0(x) + \cdots + \gamma_{g-1}(x) \cdot n^{-(g-1)})$$

appears as composed of terms

$$n^{-g}(\times l_1 \times \cdots \times l_{s-1} \times L_s \times l^*_{s+1} \times \cdots \times l^*_r) \tag{20}$$

and

$$n^{-i}(\times l_1 \times \cdots \times l_{s-1} \times \Lambda_s \times l^*_{s+1} \times \cdots \times l^*_r). \tag{21}$$

The sum of the orders of $l_1, \cdots, l_{s-1}$ is any $i < g$ and the order of L_s in (20) is $g-i$ whereas the order of Λ_s in (21) is any number $< g-i$.

---

[4] It would be the case, though, if we replaced (2) by conditions $\beta'_i N \leqq p_i \leqq \beta_i N$ with *positive* β'_i. If the differences $\beta_i - \beta'_i$ are small, this amounts to an approximate fixation of the ratio $p_1 : p_2 : \cdots : p_r$. Cf. van der Corput, l. c.[1].

We now reap our reward for having raised the lower limit in the integral J from Vinogradoff's 2 to $\sqrt{N}$: l^* and Λ vanish to the left of the point $N^{-\frac{1}{2}}$ and for $x \geqq N^{-\frac{1}{2}}$ we have $n/(n-\xi) \leqq 2$. Consequently only these two facts have to be established:

$$\times \mid l_1 \mid \times \cdots \times \mid l_r \mid \leqq C \tag{22}$$

and

$$\times \mid l_1 \mid \times \cdots \times \mid l_{s-1} \mid \times \mid \Lambda_s \mid \times \mid l_{s+1} \mid \times \cdots \times \mid l_r \mid \leqq C \cdot N^{-\frac{1}{2}}. \tag{23}$$

($l_{s+1}, \cdots, l_r$ are l's of order 0; but it is superfluous to make use of this particular circumstance.) The first inequality holds for $r \geqq 2$. If it is true for $r = 2$ it follows by induction from Lemma 1 as soon as we make sure that

$$\int \mid l(x) \mid dx \leqq C.$$

But this inequality is trivial in view of the explicit expression

$$\int_0^x (\log x)^g dx = x\{(\log x)^g - g(\log x)^{g-1} + \cdots \pm g!\}. \tag{24}$$

For $r = 2$ we have to show that

$$\int_0^x \mid (\log y/a_1)^g \cdot (\log (x-y)/a_2)^h \mid dy$$

is bounded as long as the parameter $x > 0$ stays bounded. The substitution $y = x \cdot \eta$ carries our integral over into a form

$$x \cdot \int_0^1 \mid (\log x/a_1 + \log \eta)^g \cdot (\log x/a_2 + \log (1-\eta))^h \mid d\eta$$

from which this fact follows immediately. (The essential point is that the integral tends to 0 with $x \to 0$.)

In (23) we push the factor Λ_s to the rear and apply the result (22) to $r-1$ instead of r. Formula (24) gives

$$\int \mid \Lambda \mid dx \leqq n^{-g} \cdot \int_0^{N^{-\frac{1}{2}}} \mid \log x \mid^g dx \leqq C \cdot N^{-\frac{1}{2}},$$

and hence by Lemma 1 the inequality (23) is established.

This ends our proof of Theorem IV. Moreover we learn from the inequality (22) that the coefficients $\gamma_i(x)$ are bounded functions of x; it would be quite easy to prove their continuity.

4. The singular series. The investigation of $\mathfrak{S}$ follows the familiar Hardy-Littlewood pattern. Since $\chi(q)$ has the multiplicative property $\chi(q \cdot q') = \chi(q) \cdot \chi(q')$ for any co-prime q, q', the sum $\mathfrak{S}$ is the product over all primes p of

$$1 + \chi(p) + \chi(p^2) + \cdots.$$

A simple computation yields

THEOREM V. *If d is the greatest common divisor of the numbers*

$$k;\ a_1, \cdots, a_r \tag{25}$$

then

$$\mathfrak{S}(k; a_1, \cdots a_r) = d \cdot \mathfrak{S}(k/d; a_1/d, \cdots, a_r/d). \tag{26}$$

Assuming (25) *without common divisor, we distribute the prime numbers p over $r+1$ classes, p being in the class Γ_t $(t = 0, 1, \cdots, r)$ if $t+1$ of the numbers* (25) *are prime to p. Then*

$$\mathfrak{S} = \prod_{t=0}^{r} \prod_{p \in \Gamma_t} \left\{ 1 - \frac{(-1)^t}{(p-1)^t} \right\}. \tag{27}$$

A formula like (26) was to be expected because the equation (1) coincides with

$$(a_1/d) \cdot p_1 + \cdots + (a_r/d) \cdot p_r = k/d.$$

It was also to be expected, as (27) indeed proves, that $\mathfrak{S}$ vanishes if there are prime numbers p which go into r of the numbers (25) (but not into all of them), or if

$$a_1 + \cdots + a_r \not\equiv k \pmod 2.$$

In all other cases we deduce from (27):

$$\mathfrak{S} \geqq 2 \cdot \prod_{\nu=1}^{\infty} \{1 - 1/4\nu^2\} = 4/\pi.$$

5. Prime powers. For any given integral exponent $l \geqq 1$ we may study this set of relations

$$\begin{gathered} a_1 x_1{}^l + \cdots + a_r x_r{}^l = k, \\ x_1{}^l \leqq \beta_1 N, \cdots, x_r{}^l \leqq \beta_r N. \end{gathered} \tag{28}$$

and may ask for solutions in non-negative *integers*, or more specfically in positive *primes* x_i. Let the numbers of solutions of these Waring's and Vinogradoff's problems be $A^0(N, k)$ and $A(N, k)$ respectively. If necessary we add the number r of terms as a subscript, or even give the full signature

$$\left(r_l;\ \begin{matrix} a_1, \cdots, a_r \\ \beta_1, \cdots, \beta_r \end{matrix} \right)$$

Put

$$\sigma_0(a/q) = (1/q) \cdot \sum_t e(at^l/q), \qquad \sigma(a/q) = (1/\phi(q)) \cdot \sum_t e(at^l/q)$$

where the first sum extends over *all* residues t mod q while the second sum is restricted to the reduced residues t mod q. Neither sum changes if one replaces the fraction a/q by its reduced form a'/q'. Hence we may well say that they depend on the fraction rather than on a and q separately. The singular series

$$\mathfrak{S}_0(k) = \sum_q \chi_0(q) \qquad \text{and} \qquad \mathfrak{S}(k) = \sum_q \chi(q)$$

in Waring's and Vinogradoff's problems arise from these "characters"

$$\chi_0(q) = \sum_{m \in G_q} \sigma_0(a_1 m/q) \cdot \cdot \cdot \sigma_0(a_r m/q) \cdot e(-k/q),$$

$$\chi(q) = \sum_{m \in G_q} \sigma(a_1 m/q) \cdot \cdot \cdot \sigma(a_r m/q) \cdot e(-k/q),$$

respectively. The quantitative factors R_0 and R in the asymptotic formulas

$$(29) \qquad A^0(N,k) \sim \mathfrak{S}_0(k) R_0(N,k), \qquad A(N,k) \sim \mathfrak{S}(k) R(N,k)$$

are

$$l^{-r} \int (x_1 \cdot \cdot \cdot x_r)^{\lambda-1} d\sigma \qquad \text{and} \qquad \int \frac{(x_1 \cdot \cdot \cdot x_r)^{\lambda-1} d\sigma}{\log x_1 \cdot \cdot \cdot \log x_r} \qquad (\lambda = 1/l)$$

respectively. The first integral extends over the intersection of $E(x)$ with the block

$$0 \leqq x_1 \leqq \beta_1 N, \quad \cdot \cdot \cdot, \quad 0 \leqq x_r \leqq \beta_r N,$$

the second over the intersection of the same plane with the block

$$\sqrt{N} \leqq x_1 \leqq \beta_1 N, \quad \cdot \cdot \cdot, \quad \sqrt{N} \leqq x_r \leqq \beta_r N.$$

The first integral equals $N^{\lambda r-1} \cdot R^0(k/N)$ where $R^0(x)$ is the same integral extending over $E'(x)$. (In computing these expressions we have made the substitution $x_i^l \rightarrow x_i$. The asymptotic development of $R(N,x)$ may be obtained in the same way as for $l = 1$.) We think the parallel is fairly instructive: the difference in the singular series is due to the fact that the residues of primes mod q are necessarily relatively prime to q (apart from the "few" primes going into q) whereas integers leave all residues mod q. The difference in the quantitative factor comes from the fact that 1 is the density with which the integers x are distributed over the x-axis while $1/\log x$ is the density of the primes (a fact which has to be combined with the *uniform* equidistribution over the respective residues modulo any integer q).

More irregular problems that can be tackled at all will conform to the same scheme.

The main terms of the asymptotic formulas (29) are of the order of magnitude $N^{\lambda r-1}$ and $N^{\lambda r-1}/n^r$ respectively. Of course we may claim their validity only for a sufficiently large number r of terms. The limit is set by the simpler of the two problems. Indeed, the efficiency of the method depends on two circumstances: 1) the contribution to (7) of part II of the circumference $0 \leqq \eta \leqq 1$ must be negligible; 2) the singular series must converge sufficiently well. The first is the main point. For our present prime number problem we have to set

$$S_i(\eta) = \sum_p e(a_i \eta p^l) \qquad (p^l \leqq \beta_i N).$$

Suppose that by analyzing Waring's problem we have ascertained an even number $s = 2u \geqq 4$ such that

$$A_s{}^0(N, 0) \leqq C \cdot N^{\lambda s - 1} n^K, \tag{30}$$

the integral exponent $K \geqq 0$ being independent of N. This is somewhat less than the assumption of the asymptotic formula of Waring's problem for $r = s$. It is then maintained that the asymptotic formula for Vinogradoff's problem holds good for $r \geqq 1 + s$, in the sense that the relative error ρ is of the type mentioned in Theorem I. Following Vinogradoff[5] we use the relation

$$\int_0^1 |S_1(\eta) \cdots S_u(\eta)|^2 \, d\eta = A_{2u}(N, 0) \leqq A_{2u}{}^0(N, 0)$$

where A refers to the signature

$$\left(s = 2u;\, \begin{matrix} a_1, \cdots, a_u, & -a_1, \cdots, -a_u \\ \beta_1, \cdots, \beta_u, & \beta_1, \cdots, \beta_u \end{matrix}\right),$$

and then by Schwarz's inequality and (30) we find the relation

$$\int_0^1 |S_1(\eta) \cdots S_u(\eta)| \cdot |S_{u+1}(\eta) \cdots S_{2u}(\eta)| \, d\eta \leqq C \cdot N^{\lambda s - 1} n^K$$

which forms the base for settling the first point

The accessory point, namely convergence of the singular series, is taken care of by the following estimate for σ due to Loo-keng Hua[6]:

$$|\sigma(a/q)| \leqq C_\epsilon \cdot q^{-\frac{1}{2}+\epsilon} \tag{31}$$

for $(a, q) = 1$ where ϵ is any positive number $< \frac{1}{2}$ and C_ϵ depends neither on a nor q. The requirement of "sufficiently good" convergence of the series $\Sigma\, \chi(q)$ means that it is majorized by a series $C \cdot \Sigma q^{-\alpha}$ with some exponent $\alpha > 1$. Because of (31) this is obviously the case as soon as $r \geqq 5$.

UNIVERSITY OF SASKATCHEWAN, SASKATOON, SASKATCHEWAN,
AND
INSTITUTE FOR ADVANCED STUDY, PRINCETON, N. J.

---

[5] See Vinogradoff, l.c.[2].

[6] *Math. Zeitschr.*, vol. 44 (1938), pp. 335-346, Lemma 1. 3.

LOCALLY CONNECTED SPACES AND GENERALIZED MANIFOLDS.*

EDWARD G. BEGLE.

In the course of his investigations into LC spaces (spaces locally connected in the sense of homotopy), Lefschetz introduced the fruitful notions of a realization and a partial realization of a finite complex on a space [12].[1] The chief purpose of this paper is to give new definitions of realizations and partial realizations in such a way that the methods developed for LC spaces can be carried over to lc spaces. Since there are many definitions of cycles defined for abstract spaces, there are many definitions of homology local connectedness. We choose what appears to be the weakest type of cycle, the finite Vietoris cycles. We state in Section 1 a definition of local connectedness in terms of these cycles, due originally to Alexandroff [3], and give a new definition of realizations. We find it necessary to restrict the coefficient group to be a ring with a unit. We show in Section 2 that the proofs of two important theorems carry over easily from the homotopy case. Using similar methods, we compare in Section 3 our definition of local connectedness with various others and show that they are equivalent in the important cases. We show in Section 5 that some of our results extend to compact spaces. (We follow A. Weil in using compact rather than bicompact to mean that any open covering of the space contains a finite covering. All our compact spaces are assumed to be normal. In place of compact metric space we use the convenient term compactum, introduced by Alexandroff and Hopf).

One of the more important applications of local connectedness is to generalized manifolds. These have been studied by Čech [5], Lefschetz [11], Wilder [20], and others. Since these papers appeared, there have been a number of advances in the study of homology theories of abstract spaces, and it is to be expected that some of these would permit a simplification of the discussion of these generalized manifolds. This is indeed the case. We show in Section 6 that one of the local invariants used in the definitions of these spaces, namely the local Betti number, can be phrased in terms of cocycles. Using this, we give in Section 7 a new definition of a generalized manifold and give a rapid proof of the Poincaré duality theorem for it, using cocycles and our previous results on local connectedness in an essential fashion. For

---

* Received January 23, 1941.

[1] Numbers in square brackets refer to the bibliography.

convenience we restrict ourselves to the case of an absolute, or compact, manifold. We show in Section 8 that our definition is equivalent to Čech's. We conclude by giving in Section 9 a number of examples.

The investigations undertaken in this paper were suggested by Prof. Lefschetz. We are indebted to him for his constant assistance and encouragement.

1. Definitions. In this and the next two sections the space $\mathfrak{S}$ with which we deal is a compactum. With the points of $\mathfrak{S}$ as vertices, a simplicial complex, which we denote by S, is constructed by letting any $n+1$ points of $\mathfrak{S}$, distinct or not, be the vertices of an n-simplex. Given any $\epsilon > 0$, $S[\epsilon]$ denotes the subcomplex of S consisting of all simplexes of diameter $< \epsilon$. Only finite chains on S, with coefficients taken from an untopologized abelian ring G with a unit, are considered. With such a coefficient group, a simplex can be considered as a chain. If y^q is a chain of dimension q on S, then $| y^q |$ denotes the set of q-simplexes of S on which y^q has non-zero coefficients plus all their faces. This is clearly a finite simplicial complex of dimension $\leq q$. The diameter of a chain y^q is defined to be the diameter of the set of points which are the vertices of the complex $| y^q |$.

As a matter of convenience, we make the convention that in any definition of local connectedness a 0-cycle is understood to be a 0-chain y^0 such that $KI(y^0) = 0$, where $KI(y^0)$ is the sum of the coefficients of y^0.

Definition 1.1. $\mathfrak{S}$ is said to be q-lc if for each $\epsilon > 0$ there exists a positive $\eta = \eta^q(\epsilon)$ and for each $\beta > 0$ there exists a positive $\alpha = \alpha^q(\beta, \epsilon)$ such that if x^q is a q-cycle on $S[\alpha]$ of diameter $< \eta$, then $x^q = Fy^{q+1}$, where y^{q+1} is a chain on $S[\beta]$ of diameter $< \epsilon$. $\mathfrak{S}$ is lc^p if it is q-lc for each $q \leq p$. It is lc$^\infty$ if it is q-lc for all q.

It is important to notice that η depends only on ϵ, while α depends on both β and ϵ.

Remarks I. A simple argument shows that if $\mathfrak{S}$ is lc^0 it has at most a finite number of components. If each of these is q-lc, then $\mathfrak{S}$ itself is q-lc. For if we choose $\eta^q(\epsilon)$ to be less than the minimum distance between any two components, then any cycle of diameter $< \eta^q(\epsilon)$ is on just one component. Since we are interested in $\mathfrak{S}$ only when it is at least lc^0, we can, and do, assume without any loss of generality that $\mathfrak{S}$ is connected.

II. We do not find it necessary to consider anything more general than simplicial complexes, so henceforth all complexes are understood to be simplicial. Also, with the exception of S and its subcomplexes $S[\epsilon]$, all complexes are understood to be finite.

Definition 1. 2. A ***realization*** on $S[\alpha]$ of a complex K is a chain-mapping τ defined on K with values on $S[\alpha]$, i. e., a function which assigns to each chain Y^q on K a finite chain $\tau(Y^q)$ on $S[\alpha]$ in such a way that

1) $\tau(g_1Y_1^q + g_2Y_2^q) = g_1\tau Y_1^q + g_2\tau Y_2^q, \qquad g_1, g_2 \in G;$
2) $\tau F Y^q = F\tau Y^q$
3) $KI(\tau Y^0) = KI(Y^0)$

In other words, τ is a distributive mapping which commutes with F and preserves the Kronecker index, KI, of 0-chains. The norm of τ is defined to be max [diameter $\tau(Z^q)$] for all simplexes Z^q of K.

Definition 1. 3. If K' is a subcomplex of K which contains all the vertices of K, then a realization τ' of K' on $S[\alpha]$ is said to be a partial realization of K on $S[\alpha]$. Let Z^q be a simplex of K and $Z_i^r, \cdots, Z_j^p$ all its faces which appear in K'. Then if for all Z^q and all q, max [diameter $\{\tau'(Z_i^r) + \cdots + \tau'(Z_j^p)\}$] $< \epsilon$, τ' is said to be of norm $< \epsilon$ as a partial realization of K.

Definition 1. 4. If τ' is a partial realization and τ a realization of K, we say that τ is an *extension* of τ' if τ agrees with τ' wherever the latter is defined. We also say that the partial realization τ' can be extended to the realization τ.

THEOREM 1. 1. *A necessary and sufficient condition for* $\mathfrak{S}$ *to be* lc^p *is that for each* $\epsilon > 0$ *there exist a positive* $\bar{\eta} = \bar{\eta}^p(\epsilon)$ *and for each* $\beta > 0$ *there exist a positive* $\bar{\alpha} = \bar{\alpha}^p(\beta, \epsilon)$ *such that if* τ' *is a partial realization on* $S[\bar{\alpha}]$ *of norm* $< \bar{\eta}$ *of a complex* K *of dimension* $\leq p + 1$, *then* τ' *can be extended to a realization* τ *of* K *on* $S[\beta]$ *of norm* $< \epsilon$.

Proof. For the sufficiency let $\eta^q(\epsilon) = \frac{1}{2}\bar{\eta}^p(\epsilon/3)$ and $\alpha^q(\beta, \epsilon) = \bar{\alpha}^p(\beta, \epsilon/3)$ for any $q \leq p$. Let x^q be a cycle on $S[\bar{\alpha}]$ of diameter $< \frac{1}{2}\bar{\eta}$. Let K be the join of $|x^q|$ and one of its vertices, z^0, and let τ' be the identity mapping on $|x^q|$. Then τ' is a partial realization of K on $S[\bar{\alpha}]$ of norm $< \bar{\eta}$. Hence τ' can be extended to a realization τ of K on $S[\beta]$ of norm $< \epsilon/3$. Let y^{q+1} be the join of x^q and z^0 so that $Fy^{q+1} = x^q$. Let $y_1^{q+1} = \tau(y^{q+1})$. Then y_1^{q+1} is on $S[\beta]$ and is of diameter $< 2\epsilon/3 + \bar{\eta} < \epsilon$. Since τ is an extension of τ' and τ' is the identity on $|x^q|$, $Fy_1^{q+1} = F\tau(y^{q+1}) = \tau(Fy^{q+1}) = \tau(x^q) = x^q$, which completes the proof of the sufficiency.

The necessity being true for $p = 0$, because of condition 3) in definition 1. 2, we may assume the theorem proved for $p - 1$, so that we have the functions $\eta^p(\epsilon)$, $\alpha^p(\beta, \epsilon)$, $\bar{\eta}^{p-1}(\epsilon)$, and $\bar{\alpha}^{p-1}(\beta, \epsilon)$. Choose $\gamma = (1/3)\eta^p(\epsilon)$ and $\lambda = \alpha^p(\beta, \epsilon)$, and set $\bar{\eta}^p(\epsilon) = \bar{\eta}^{p-1}(\gamma)$ and $\bar{\alpha}^p(\beta, \epsilon) = \bar{\alpha}^{p-1}(\lambda, \gamma)$.

Now let K be a complex of dimension $\leq p + 1$, and τ' a partial realization of K on $S[\bar{\alpha}^p]$ of norm $< \bar{\eta}^p$. Denote by K^p the p-section of K, i. e., the set of simplexes of K of dimension $\leq p$. Let τ'' be τ' considered only on K^p. Then τ'' is a partial realization of K^p on $S[\bar{\alpha}^p]$ of norm $< \bar{\eta}^p$. By the choice of $\bar{\eta}^p$ and $\bar{\alpha}^p$, τ'' can be extended to a realization τ_1 of K^p on $S[\lambda]$ and of norm $< \gamma$. Let τ_2 be the mapping which is τ' wherever the latter is defined and is τ_1 otherwise, and let Z^{p+1} be a simplex of K which is not mapped by τ_2. FZ^{p+1} is mapped by τ_2 onto a cycle x^p on $S[\lambda]$, and diameter $x^p < \bar{\eta}^{p-1}(\gamma) + 2\gamma < 3\gamma < \eta^p(\epsilon)$. Hence x^p bounds a chain y^{p+1} on $S[\beta]$ such that diameter $y^{p+1} < \epsilon$. Define $\tau(Z^{p+1})$ to be y^{p+1}, $\tau(\sum_i g_i Z_i^{p+1})$ to be $\sum_i g_i \tau(Z_i^{p+1})$, and τ to be τ_2 otherwise. Then τ is an extension of τ' and is defined on all of K, which completes the proof.

The functions $\bar{\eta}^p(\epsilon)$ and $\bar{\alpha}^p(\beta, \epsilon)$ are called *extension functions.*

Definition 1. 5. $\mathfrak{S}$ is lc if for each $\epsilon > 0$ there exists a positive $\eta = \eta(\epsilon)$ and for each $\beta > 0$ there exists a positive $\alpha = \alpha(\beta, \epsilon)$ such that if τ' is a partial realization on $S[\alpha]$ of norm $< \eta$ of a complex K, then τ' can be extended to a realization τ of K on $S[\beta]$ of norm $< \epsilon$.

This definition is stronger than that of lc$^\infty$. However, we have

THEOREM 1. 2. *If* $\mathfrak{S}$ *is* lc^n *and dimension*[2] $\mathfrak{S} \leq n$, *then* $\mathfrak{S}$ *is* lc.

The proof is postponed to Section **5**.

2. Properties of lc^n spaces. Two important properties of locally connected spaces can now be proved in essentially the same way as they are for the homotopy local connectedness case. It is understood that the homolgy groups, $H^p(\mathfrak{S}, G)$ and the Betti numbers for compacta are defined in terms of Vietoris cycles [18].

THEOREM 2. 1. *If* $\mathfrak{S}$ *is* lc^n, *then its homology groups* $H^p(\mathfrak{S}, G)$, *for* $p \leq n$, *are isomorphic with subgroups of the homology groups of a finite complex.*[3]

This theorem is a corollary of the more general theorem 5. 5, proved later.

COROLLARY 2. 2. *If* $\mathfrak{S}$ *is* lc^n, *its p-dimensional Betti numbers, for* $p \leq n$, *are all finite.*

THEOREM 2. 3. *The Lefschetz fixed point formula is valid for* lc *spaces.*

The proof is essentially the same as that of Lefschetz [13], so we omit it.

Remark. Define $\mathfrak{S}$ to be weak-lc^n if it is lc^{n-1} and n-lc except that the chain y^{n+1} of definition 1.1 need not be of diameter $< \epsilon$. Then it is clear that Theorems 2.1, 2.2, and 2.3 are true if $\mathfrak{S}$ is weak-lc^n [21].

---

[2] Menger-Urysohn dimension.

[3] This theorem was proved by Alexandroff in [3] for the case of integer coefficients in the definition of local connectedness. See also [21].

In order to establish a connection between the type of local connectedness which we have defined and the ones in terms of Vietoris cycles and projection cycles, we prove the following lemma. We recall first that a projection chain[4] Δ^p is a sequence $\{y_i^p\}$ of chains of S with the following properties: 1) y_i^p is on $S[\epsilon_i]$ and $\epsilon_i \to 0$, 2) for each i there is defined a simplicial projection π_i such that $\pi_i y_i^p = y_{i-1}^p$, and 3) denoting the compound projection $\pi_{j+1}\pi_{j+2} \cdots \pi_i$ by π_i^j, for each $\epsilon > 0$ there is an n such that for $i > j > n$, π_i^j is an ϵ-projection, i. e., no vertex of y_i^p is projected a distance greater than ϵ.

LEMMA 2.4. *If $\mathfrak{S}$ is* lc^{n-1}, *then for each $\mu > 0$ there is a $\lambda > 0$ such that if y^p, $p \leq n$, is a chain on $S[\lambda]$, then it is the first element of a projection chain Δ^p such that diameter $\Delta^p < 2\mu +$ diameter y^p and such that for all i, π_i^1 is a λ-projection.*

Proof. Given $\mu > 0$, choose a sequence $\{\mu_i\}$ such that each $\mu_i > 0$, and $\sum_1^\infty \mu_i < \mu$. Then $\lambda = \bar{\eta}^{n-1}(\mu_1)$ satisfies the conditions of the lemma. To demonstrate this, first construct an auxiliary sequence $\{\nu_i\}$, $\nu_i > 0$ such that $\sum_1^\infty \nu_i < \lambda$ and also such that $\nu_i < \bar{\eta}^{n-1}(\mu_{i+1})$.

Now let y^p be a chain on $S[\lambda]$. The identity mapping on the vertices of $| y^p |$ is a partial realization τ'_1 of $| y^p |$ on $S[\bar{\alpha}^{n-1}(\nu_1, \mu_1)]$ and τ'_1 is of norm $< \lambda = \bar{\eta}^{n-1}(\mu_1)$. Hence τ'_1 can be extended to a realization τ_1 of $| y^p |$ on $S[\nu_1]$ and of norm $< \mu_1$.

Let z^q and $\bar{z}^q$ be simplexes of $| y^p |$. If $| \tau_1(z^q) |$ and $| \tau_1(\bar{z}^q) |$ have common points, we wish these to be in a $| \tau_1(z^r) |$, where z^r is a common face of z^q and $\bar{z}^q$. If this is not the case, we replace the simplexes of $\tau(z^q)$ by nearby simplexes so as to effect this. Since $\mathfrak{S}$ is connected, there are infinitely many points of $\mathfrak{S}$ arbitrarily near any given point, and hence this shifting can be done in such a way as to keep norm $\tau_1 < \mu_1$. The new mapping we continue to call τ_1. Without further mention the mappings τ_i now to be constructed are modified, if necessary, in the same way. The reason for these modifications will appear presently.

The chain $y_2^p = \tau_1(y^p)$ is now on $S[\nu_1]$ and its diameter is $< 2\mu_1 +$ diameter y^p. We repeat the above process on y_2^p. The identity mapping on the vertices of $|y_2^p|$ forms a partial realization τ'_2 of $|y_2^p|$ on $S[\bar{\alpha}^{n-1}(\nu_2, \mu_2)]$ and norm $\tau'_2 < \nu_1 < \bar{\eta}^{n-1}(\mu_2)$. Hence τ'_2 can be extended to a realization τ_2 of $|y_2^p|$ of norm $< \mu_2$ on $S[\nu_2]$. $\tau_2(y_2^p)$ is therefore a chain y_3^p on $S[\nu_2]$ and diameter $y_3^p < 2\mu_2 +$ diameter $y_2^p < 2\mu_2 + 2\mu_1 +$ diameter y^p.

---

[4] These chains have been studied by Alexandroff [3] and Lefschetz [14]. The projection cycles do not form a group and do not give rise to a homology group. In order to obtain a homology group, it is necessary to use $\mathfrak{P}$-chains [14], which are classes of projection chains.

Proceeding in this fashion we arrive at a sequence $\Delta^p = \{y_i^p\}$, $y_1^p = y^p$, where y_i^p is on $S[\nu_{i-1}]$ and $\nu_i \to 0$. The diameter of Δ^p is $<$ diameter $y^p + 2\sum_1^\infty \mu_i <$ diameter $y^p + 2\mu$.

To show that Δ^p is actually a projection-chain, we must define projections π_i such that $\pi_{i+1}(y_{i+1}^p) = y_i^p$. Starting with the simplexes of dimension 1, for any z^1 of $|y_i^p|$, we map each vertex of $|\tau_i(z^1)|$ into the nearest vertex of z^1. Then for any z^2 of $|y_i^p|$, we map into the nearest vertex of z^2 those vertices of $|\tau_i(z^2)|$ which have not already been mapped. Proceeding in this way, after p steps we have mapped all of $|\tau_i(y_i^p)|$ into $|y_i^p|$. Because of the modifications made in each τ_i, there is no ambiguity in this definition. This mapping is π_{i+1}. The amount of π_{i+1} is clearly $< \nu_i$, so that for any i, the amount of $\pi_i^1 = \pi_2 \cdots \pi_{i-1}\pi_i$ is $< \sum_1^i \nu_k < \sum_1^\infty \nu_k < \lambda$.

To see that $\pi_{i+1}(y_{i+1}^p) = y_i^p$, we note first that the sub-chain of y_{i+1}^p which maps into z^p is $\tau_i(z^p)$. Suppose $\pi_{i+1}\tau_i(z^p) = gz^p$. If the vertices of z^p are $A_0, \cdots, A_p$, denote by Z_k the simplex whose vertices are $A_k, \cdots, A_p$. Since τ_i and π_{i+1} both commute with F for chains over G, we have,

$$gZ_1 + \cdots = Fgz^p = F\pi_{i+1}\tau_i(z^p) = \pi_{i+1}\tau_i(Z_1) + \cdots$$

Hence

$$gZ_1 = \pi_{i+1}\tau_i(Z_1)$$

and

$$\begin{aligned} gZ_2 + \cdots &= FgZ_1 = F\pi_{i+1}\tau_i(Z_1) = \pi_{i+1}\tau_i(Z_2) + \cdots \\ gZ_3 + \cdots &= FgZ_2 = F\pi_{i+1}\tau_i(Z_2) = \pi_{i+1}\tau_i(Z_3) + \cdots \\ &\;\;\vdots \\ gA_p + \cdots &= FgZ_{p-1} = F\pi_{i+1}\tau_i(Z_{p-1}) = \pi_{i+1}\tau_i(A_p). \end{aligned}$$

But

$$gA_p = \pi_{i+1}\tau_i(A_p) = A_p,$$

since π_{i+1} and τ_i are the identity on the vertices of $|y_i^p|$. Therefore $g = 1$, and $\pi_{i+1}(y_{i+1}^p) = y_i^p$, which completes the proof

3. Equivalence theorem. By means of Lemma 2.4, we can show that our definition of local connectedness is equivalent to the usual one in terms of

Vietoris cycles and also to the one in terms of projection cycles. These are defined as follows:

Definition 3.1. $\mathfrak{S}$ is said to be $\mathrm{lc}_V{}^n[\mathrm{lc}_P{}^n]$ if for each $\epsilon > 0$ there is a positive $\eta = \eta_V{}^n(\epsilon)\,[\eta_P{}^n(\epsilon)]$ such that a Vietoris [projection] cycle of dimension $\leq n$ and of diameter $< \eta_V{}^n[\eta_P{}^n]$ is ~ 0 in a set of diameter $< \epsilon$ [bounds a projection chain of diameter $< \epsilon$].[5]

Denote the class of all compacta which are lc^n or locally connected in dimensions $\leq n$ with respect to Vietoris or projection cycles by $\{\mathrm{lc}^n\}$, $\{\mathrm{lc}_V{}^n\}$, and $\{\mathrm{lc}_P{}^n\}$ respectively.

THEOREM 3.1. $\{\mathrm{lc}^n\} = \{\mathrm{lc}_V{}^n\} = \{\mathrm{lc}_P{}^n\}$.

Proof. a) $\{\mathrm{lc}^n\} \subset \{\mathrm{lc}_V{}^n\}$.

Let $\Gamma^n = \{x_i{}^n\}$ be a Vietoris cycle of diameter $< \eta^n(\epsilon/3)$. For any $\beta > 0$ there exists a number k_β such that for $i > k_\beta$, $x_i{}^n$ is on $S[\alpha^n(\beta, \epsilon/3)]$. Then $x_i{}^n = Fy^{n+1}$, where y^{n+1} is of diameter $< \epsilon/3$ and is on $S[\beta]$. Hence $\Gamma^n \sim 0$ is a set of diameter $< 2\epsilon/3 + \eta^n(\epsilon/3) < \epsilon$, which proves the inclusion.

b) $\{\mathrm{lc}_V{}^n\} \subset \{\mathrm{lc}^n\}$

This is proved by induction. It is obviously true for $n = 0$, so we may assume it true for $n - 1$. Hence $\mathfrak{S}$ is lc^{n-1} and $\mathrm{lc}_V{}^n$. Given ϵ, $\beta > 0$, we let $\eta^n(\epsilon) = \eta_V{}^n(\epsilon/3)/3$. Let $\mu < \eta^n(\epsilon)$ and take $\lambda < \mu$ by Lemma 2.4, and $< \beta/3$. Set $\alpha^n(\beta, \epsilon) = \lambda$.

To show that $\eta^n(\epsilon)$ and $\alpha^n(\beta, \epsilon)$ are satisfactory, let x^n be of diameter $< \eta^n(\epsilon)$, and let it be on $S[\alpha^n]$. Then by Lemma 2.4 there exists a projection cycle $\Delta^n = \{x_i{}^n\}$, $x_1{}^n = x^n$, of diameter $< \eta_V{}^n(\epsilon/3)/3 + 2\mu < \eta_V{}^n(\epsilon/3)$. Δ^n, considered as a Vietoris cycle, is therefore ~ 0 in a set of diameter $< \epsilon/3$. Hence, for i sufficiently large, $x_i{}^n = Fy_1{}^{n+1}$, where $y_i{}^{n+1}$ is on $S[\beta/3]$, and is of diameter $< \epsilon/3$.

Now let δ_i be the deformation chain[6] associated with $\pi_i{}^1$, so that $F\delta_i x_i{}^n = x_1{}^n - x_i{}^n$. Then $y^{n+1} = \delta_i x_i{}^n + y_i{}^{n+1}$ is on $S[\beta]$ since $\lambda < \beta/3$ and $\pi_i{}^1$ is a λ-transformation. Also, diameter $y^{n+1} < \epsilon/3 + 2\lambda < \epsilon$. Finally, $Fy^{n+1} = x_1{}^n = x^n$, which completes the proof of the inclusion b).

c) $\{\mathrm{lc}_P{}^n\} \subset \{\mathrm{lc}^n\}$.

The proof of b) applies to this case also.

d) $\{\mathrm{lc}^n\} \subset \{\mathrm{lc}_P{}^n\}$.

---

[5] It is not hard to see that local connectedness with respect to $\mathfrak{P}$-chains is equivalent to local connectedness with respect to projection chains.

[6] For the delnition and properties of such chains, see [14].

Set $\eta_P^n(\epsilon) = \eta^n(\epsilon/3)$ and let $\Delta^n = \{x_i^n\}$ be a projection cycle of diameter $< \eta_P^n(\epsilon)$. It is sufficient to show that a subsequence of Δ^n is the boundary of a projection chain. Hence we may assume that x_1^n is on $S[\alpha^n(\lambda, \epsilon/3)]$, where λ is given by Lemma 2.4 for $\mu = \epsilon/3$. Then $x_1^n = Fy_1^{n+1}$, where y_1^{n+1} is on $S[\lambda]$ and is of diameter $< \epsilon/3$. The procedure used in the proof of Lemma 2.4 can be applied with τ'_i now being the identity mapping on $| x^n_{i+1} |$ plus the vertices of $| \delta_{i+1}(x^n_{i+1}) + y_i^{n+1} |$ not in $| x^n_{i+1} |$, δ_{i+1} being the deformation associated with π_{i+1}. In this way a projection chain Δ^{n+1} of diameter $< \epsilon$ is constructed such that $F\Delta^{n+1} = \Delta^n$. The proof of this differs only in details from that of Lemma 2.4, so it is omitted.

4. Extension to compact spaces. The definition of q-lc is in terms of finite chains and cycles defined on the infinite complexes $S[\epsilon]$. Paralleling this there is a definition due to Čech [8] in terms of chains and cycles defined on finite complexes, namely the nerves of coverings of $\mathfrak{S}$.

Since the homology theory of non-metric spaces is usually in terms of Čech cycles, this latter definition is more suitable for any consideration of spaces more general than compacta. Below we rephrase our definition of q-lc in terms of the open sets of the space, instead of the metric, so that it has meaning for any space and then show that if the space is compact, it is equivalent to Čech's.

The notation used is as follows: All coverings of $\mathfrak{S}$ are understood to be finite. Open coverings of $\mathfrak{S}$ are denoted by $\mathfrak{A}, \mathfrak{B}, \cdots$, and their nerves by $A, B, \cdots$. The covering consisting of the closures of the open sets of $\mathfrak{A}$ is denoted by $\mathfrak{A}^o$, and its nerve by A^o. If $\mathfrak{A}$ and $\mathfrak{B}$ are coverings, $\mathfrak{A} \cup \mathfrak{B}$ is the covering consisting of all the open sets of $\mathfrak{A}$ and $\mathfrak{B}$ together, and its nerve is denoted by $A \cup B$.

We write $\mathfrak{A} < \mathfrak{B}$ if $\mathfrak{A}$ is a refinement of $\mathfrak{B}$, and $\mathfrak{A} \ll \mathfrak{B}$ if $\mathfrak{A}^o$ is a refinement of $\mathfrak{B}$.

If P is a subset of $\mathfrak{S}$ and $\mathfrak{A}$ is a covering, $St\ (P; \mathfrak{A})$, the star of P in $\mathfrak{A}$, is the set which is the sum of all the sets of $\mathfrak{A}$ which meet P. If $\mathfrak{A}$ and $\mathfrak{B}$ are coverings, $St\ (\mathfrak{B}; \mathfrak{A})$ is the covering whose elements are the stars in $\mathfrak{A}$ of the elements of $\mathfrak{B}$.

We write $\mathfrak{A} * < \mathfrak{B}$ if $St\ (\mathfrak{A}; \mathfrak{A}) < \mathfrak{B}$ and say that $\mathfrak{A}$ is a star-refinement of $\mathfrak{B}$.

If $\mathfrak{A} < \mathfrak{B}$, we denote by π_a^b any one of the simplicial projections of A into B.

Homology groups of $\mathfrak{S}$ are of course defined in terms of Čech cycles. We use the Lebesgue type of dimension for $\mathfrak{S}$, i. e., $\mathfrak{S}$ is of dimension $\leq n$ if any covering has a refinement the dimension of whose nerve is no greater than n [6].

A simplex of A is said to be on a set when its kernel (the intersection of the open sets which are its vertices) meets the set. If Y^q is a chain of A, we write diameter $Y^q < \mathfrak{E}$ if there is one set of $\mathfrak{E}$ such that every simplex of $| Y^q |$ is on this set. Note that if $\mathfrak{A} < \mathfrak{B}$, and Y^q is a chain on A of diameter $< \mathfrak{E}$, then $\pi_a{}^b Y^q$ is also of diameter $< \mathfrak{E}$.

The infinite complex S is defined as in Section 1, but the subcomplexes $S[\epsilon]$ are now replaced by the subcomplexes $S[\mathfrak{A}]$, a simplex of S being in $S[\mathfrak{A}]$ if there is one set of $\mathfrak{A}$ which contains all its vertices. If y^q is a chain of S, we write diameter $y^q < \mathfrak{E}$ if there is a set of $\mathfrak{E}$ which contains all the vertices of the complex $| y^q |$.

With each covering $\mathfrak{A}$ is associated a subcomplex a of S as follows: the vertices of a are distinct points, one in each of the sets of $\mathfrak{A}$. That such distinct points exist follows from the fact that $\mathfrak{S}$ is connected. Any $n + 1$ of these points are the vertices of a simplex of a if the corresponding sets of $\mathfrak{A}$ intersect. Thus a and A are isomorphic complexes. If Y is any chain of A, there is a corresponding chain y of a, and conversely. If $\mathfrak{A}^* < \mathfrak{B}$, then a is a subcomplex of $S[\mathfrak{B}]$.

We now have two kinds of realizations of a complex K. Formally, these are defined as in definition 1.2, but in one case the image under τ of a chain on K is a chain on a nerve A, and in the other the image is a finite chain on $S[\mathfrak{A}]$. The norm of τ is defined in the obvious way, as is the norm of a partial realization of either type.

In the following, whenever we write functions $\mathfrak{A}(\mathfrak{B}, \cdots)$ it is understood that $\mathfrak{A} < \mathfrak{B}, \cdots$.

With the above notations, we can now rephrase definition 1.1 as follows:

Definition 4.1. $\mathfrak{S}$ is q-lc if, given coverings $\mathfrak{E}$ and $\mathfrak{B}$, there exist coverings $\mathfrak{H}^q = \mathfrak{H}^q(\mathfrak{E})$ and $\mathfrak{A}^q = \mathfrak{A}^q(\mathfrak{B}, \mathfrak{E})$ such that if x^q is a cycle on $S[\mathfrak{A}^q]$ of diameter $< \mathfrak{H}^q$, then $x^q = F y^{q+1}$, where y^{q+1} is a chain on $S[\mathfrak{B}]$ of diameter $< \mathfrak{E}$.

Using the fact that attached to each open covering of a compactum there is a number ν such that any set of diameter $< \nu$ is contained in a set of the covering, it is easy to show that definitions 4.1 and 1.1 are equivalent in such spaces.

For a compact space, the definition of Čech can be stated as:

Definition 4.2. $\mathfrak{S}$ is q-Clc if, given coverings $\mathfrak{E}$ and $\mathfrak{B}$, there exist coverings $\mathfrak{H}^q = \mathfrak{H}^q(\mathfrak{E})$ and $\mathfrak{A}^q = \mathfrak{A}^q(\mathfrak{B}, \mathfrak{E})$ such that if X^q is a cycle on A^q of diameter $< \mathfrak{H}^q$, then $\pi_a{}^b X^q = F Y^{q+1}$, where Y^{q+1} is a chain on B of diameter $< \mathfrak{E}$.

For some purposes it is more convenient to have this definition in the following form:

Definition 4.3. $\mathfrak{S}$ is q-Clc if, given coverings $\mathfrak{E}$ and $\mathfrak{B}$, there exist coverings $\mathfrak{H}^q = \mathfrak{H}^q(\mathfrak{E})$ and $\mathfrak{A}^q = \mathfrak{A}^q(\mathfrak{B}, \mathfrak{E})$ such that if X^q is a cycle on A^q, of diameter $< \mathfrak{H}^q$, then $X^q = FY_1^{q+1}$, where Y_1^{q+1} is a chain on $A^q \cup B$ of diameter $< \mathfrak{E}$.

To see that definitions 4.2 and 4.3 are equivalent let Y^{q+1} be $\pi_a{}^b Y_1{}^{q+1}$ and conversely let $Y_1{}^{q+1} = Y^{q+1} + \delta X^q$, where δ is the deformation associated with $\pi_a{}^b$.

It is clear that in definitions 4.2 and 4.3 $\mathfrak{H}^q$ can be replaced by any $\mathfrak{H} < \mathfrak{H}^q$, and $\mathfrak{A}^q$ can be replaced by any $\mathfrak{A} < \mathfrak{A}^q$. Also, if these definitions are satisfied for any $\mathfrak{E}$ and $\mathfrak{B}$ chosen from a complete family of coverings, then they are clearly satisfied for arbitrary $\mathfrak{E}$ and $\mathfrak{B}$.

We shall need the following facts about compact spaces.

4.a) Every covering $\mathfrak{B}$ has a star-refinement, $\mathfrak{A}$; $\mathfrak{A}^* < \mathfrak{B}$ [5, p. 625].

4.b) If $\mathfrak{B}$ is any covering, each set of $\mathfrak{B}$ contains an open subset such that these sets form a covering $\mathfrak{A}$, and $\mathfrak{A} << \mathfrak{B}$ [2, p. 73].

4.c) If $\mathfrak{A} << \mathfrak{B}$, there is a covering $\mathfrak{C}$ such that $St\,(\mathfrak{A};\mathfrak{C}) < \mathfrak{B}$.

The proof is trivial.

We say that $\mathfrak{C}$ is a Lebesgue refinement of $\mathfrak{A}$ if $\mathfrak{C} < \mathfrak{A}$ and if whenever a collection of sets of $\mathfrak{A}$ meets a set of $\mathfrak{C}$ they have a non-vacuous intersection.

4.d) Each covering $\mathfrak{B}$ has a refinement $\mathfrak{A}$ which has a Lebesgue refinement $\mathfrak{C}$.

The proof is as follows: By 4.b) there is a sequence, $\{\mathfrak{A}_i\}$, $\mathfrak{A}_0 = \mathfrak{B}$, $\mathfrak{A}_i << \mathfrak{A}_{i-1}$ such that $A_i{}^c$ is isomorphic to a subcomplex of A_{i-1}. Since B is a finite complex, there is a least integer k such that $A_k{}^c$ is isomorphic to A_k. This means that if the closures of a collection of sets of $\mathfrak{A}_k$ have a non-vacuous intersection, the same is true of the sets themselves. Let $\mathfrak{A} = \mathfrak{A}_k$.

Denote by $\{\mathfrak{U}_\zeta\}$ the collection of all refinements of $\mathfrak{A}$, and suppose that $\mathfrak{A}$ has no Lebesgue refinement. Then for some collection J of sets of $\mathfrak{A}$, and for each ζ, there is an open set M_ζ of $\mathfrak{U}_\zeta$ such that the sets of J all meet M_ζ and at the same time these sets, and their closures, have a vacuous intersection. Let s_ζ be a point of M_ζ. Then $\{s_\zeta\}$ is a directed system.[7] Since $\mathfrak{S}$ is compact, $\{s_\zeta\}$ has a cluster point s. Let N be any neighborhood of s. By the definition of cluster point, it follows that there is a ζ such that M_ζ is in N. Hence there are points of each of the sets of J in N, and so s is a limit point of each of these sets, and the closures of these sets have a non-vacuous intersection, which is a contradiction. Therefore one of the $\mathfrak{U}_\zeta$ is a Lebesgue refinement of $\mathfrak{A}$.

---

[7] For the definition of directed systems and cluster points, see [17].

5. Properties of compact lc* spaces. In this and all the following sections, $\mathfrak{S}$ is understood to be compact.

THEOREM 5. 1. *If* $\mathfrak{S}$ *is* q-lc, *it is also* q-Clc *and conversely.*

Proof. a) q-lc implies q-Clc. Denote refinements of $\mathfrak{E}$ and $\mathfrak{B}$ given by definition 4. 1 by $\mathfrak{H}_{(1)}$ and $\mathfrak{A}_{(1)}$. Given any covering $\mathfrak{E}$, there is a covering $\mathfrak{E}_1 << \mathfrak{E}$. Also, there is a covering $\mathfrak{H} << \mathfrak{H}_{(1)}{}^q(\mathfrak{E}_1)$. Given $\mathfrak{B}$, by 4. c), 4. d) there is a covering $\mathfrak{B}_1 < \mathfrak{B}$ such that $St(\mathfrak{E}_1; \mathfrak{B}_1) < \mathfrak{E}$ and such that $\mathfrak{B}_1$ has a Lebesgue refinement, $\mathfrak{B}_2$. By 4. a) there is a covering $\mathfrak{A}_1 < \mathfrak{A}^q{}_{(1)}(\mathfrak{B}_2, \mathfrak{E}_1)$ such that $St(\mathfrak{H}_1; \mathfrak{A}_1) < \mathfrak{H}^q{}_{(1)}(\mathfrak{E}_1)$. Finally there is a covering $\mathfrak{A}_2{}^* < \mathfrak{A}_1$. Now the coverings $\mathfrak{H}^q(\mathfrak{E}) = \mathfrak{H}$ and $\mathfrak{A}^q(\mathfrak{B}, \mathfrak{E}) = \mathfrak{A}_2$ satisfy the conditions of definition 4. 2).

For let X^q be a cycle of A_2 of diameter $< \mathfrak{H}$. Then x^q of a_2 is on $S[\mathfrak{A}_1]$ and is of diameter $< \mathfrak{H}^q{}_{(1)}(\mathfrak{E}_1)$. Therefore $x^q = Fy^{q+1}$, where y^{q+1} is on $S[\mathfrak{B}_2]$ and is of diameter $< \mathfrak{E}_1$. We map y^{q+1} onto a chain $Y_1{}^{q+1}$ of B_1 as follows: for each vertex of y^{q+1} which is a vertex of x^q there is an open set of $\mathfrak{A}_2$ which contains it and an open set of $\mathfrak{B}_1$ into which the latter is projected by $\pi_{a_2}{}^{b_1}$. This vertex of B_1 is chosen as the image of the vertex of x^q. Any vertex of y^{q+1} not a vertex of x^q is mapped into an open set of $\mathfrak{B}_1$ which contains it. By the choice of $\mathfrak{B}_2$, this mapping is in fact a simplicial mapping, ϕ, so $\phi(y^{q+1}) = Y_1{}^{q+1}$ is a chain of B_1, and $FY_1{}^{q+1} = \phi(x^q)$. But by construction $\phi(x^q) = \pi_{a_2}{}^{b_1} X^q$. Finally, diameter $Y_1{}^{q+1} < \mathfrak{E}$, so $\pi_{b_1}{}^b Y_1{}^{q+1} = Y^{q+1}$ is on B and is of diameter $< \mathfrak{E}$, and $FY^{q+1} = \pi_a{}^b X^q$, which completes the proof of a).

b) q-Clc implies q-lc.

Given any covering $\mathfrak{E}$, there exist $\mathfrak{E}_1 << \mathfrak{E}$ and $\mathfrak{H}_{(1)} << \mathfrak{H}^q(\mathfrak{E}_1)$. Given $\mathfrak{B}$, there is a covering $\mathfrak{B}_1 < \mathfrak{B}$ such that $St(\mathfrak{E}_1; \mathfrak{B}_1) < \mathfrak{E}$. Also there is a covering $\mathfrak{B}_2{}^* < \mathfrak{B}_1$ and a covering $\mathfrak{B}_3$ such that $\mathfrak{B}_3 < \mathfrak{A}^q(\mathfrak{B}_2; \mathfrak{H}_{(1)})$, such that $St(\mathfrak{H}_{(1)}; \mathfrak{B}_3) < \mathfrak{H}^q(\mathfrak{E}_1)$ and such that $\mathfrak{B}_3$ has a Lebesgue refinement, $\mathfrak{A}_{(1)}$. Now the coverings $\mathfrak{H}^q{}_{(1)}(\mathfrak{E}) = \mathfrak{H}_{(1)}$ and $\mathfrak{A}^q{}_{(1)}(\mathfrak{B}; \mathfrak{E}) = \mathfrak{A}_{(1)}$ satisfy the conditions of definition 4. 1.

For let x^q be a cycle of $S[\mathfrak{A}_{(1)}]$ of diameter $< \mathfrak{H}_{(1)}$. Mapping each vertex of x^q onto a vertex of B_3 which contains it defines a mapping ψ of x^q onto B_3, and since $\mathfrak{A}_{(1)}$ is a Lebesgue refinement of $\mathfrak{B}_3$, ψ is simplicial, so $\psi(x^q) = X_1{}^q$ is a cycle of B_3. Also diameter $X_1{}^q < \mathfrak{H}^q(\mathfrak{E}_1)$, so $X_1{}^q = FY_1{}^{q+1}$, by hypothesis, where $Y_1{}^{q+1}$ is on $B_3 \vee B_2$ and is of diameter $< \mathfrak{E}_1$.

To Y_1^{q+1} there corresponds a chain y_1^{q+1} of $S[\mathfrak{B}_1]$, and diameter $y_1^{q+1} < \mathfrak{E}$. Let $Fy_1^{q+1} = x_1^{q+1}$. But corresponding to the mapping ψ there is a simplicial mapping ϕ of x^q onto x_1^q. The corresponding deformation chain being δx^q and setting $y_1^{q+1} + \delta x^q = y^{q+1}$, we have $Fy^{q+1} = x^q$, where y^{q+1} is on $S[\mathfrak{B}_1] \subset S[\mathfrak{B}]$ and is of diameter $< \mathfrak{E}$, which completes the proof.

One of the important consequences of this theorem is that the large body of theorems developed by Čech for Clc^n spaces now applies, in the case of compacta, to the types of local connectedness defined in Section **3**.

By the same procedure as was used in the proof of Theorem 1. 1, the following theorem can be established.

THEOREM 5. 2. *A necessary and sufficient condition that* $\mathfrak{S}$ *be* Clc *is that for each covering* $\mathfrak{E}$ *there exist a covering* $\bar{\mathfrak{H}}^p = \bar{\mathfrak{H}}^p(\mathfrak{E})$ *and for each covering* $\mathfrak{B}$ *there exist a covering* $\bar{\mathfrak{A}}^p = \bar{\mathfrak{A}}^p(\mathfrak{B}, \mathfrak{E})$ *such that if* K *is a complex of dimension* $\leq p + 1$ *and* τ' *is a partial realization of* K *on* $\bar{A}$ *of norm* $< \bar{\mathfrak{H}}^p$, *then* τ' *can be extended to a realization* τ *of* K *on* $\bar{A} \cup \bar{B}$ *of norm* $< \mathfrak{E}$.

The functions $\bar{\mathfrak{H}}^p = \bar{\mathfrak{H}}^p(\mathfrak{E})$ and $\bar{\mathfrak{A}}^p = \bar{\mathfrak{A}}^p(\mathfrak{B}, \mathfrak{E})$ are called *extension functions.* It is clear that the statement of Theorem 5. 2 is true for any $\mathfrak{H} < \bar{\mathfrak{H}}^p$ and for any $\mathfrak{A} < \bar{\mathfrak{A}}^p$.

The definitions of lc^∞, Clc^∞, lc, and Clc are now as in definitions 1. 1 and 1. 5 with the obvious changes. By the same method as was used in the proof of Theorem 5. 1, a proof can be found for

THEOREM 5. 3. *If* $\mathfrak{S}$ *is* lc, *it is also* Clc *and conversely.*

Čech has shown that if dimension[8] $\mathfrak{S} \leq n$, then $\mathfrak{S}$ is q-Clc for all $q > n$. Hence if dimension $\mathfrak{S} \leq n$ and if $\mathfrak{S}$ is Clc^n, then it is Clc^∞. However, we can establish the stronger

THEOREM 5. 4. *If dimension* $\mathfrak{S} \leq n$ *and* $\mathfrak{S}$ *is* Clc^n, *then it is* Clc.

Proof. By Čech's result, $\mathfrak{S}$ is Clc^{2n+2}. Given $\mathfrak{E}$, let $\mathfrak{E}_1$ be $* < \mathfrak{E}$. Given any covering $\mathfrak{B}$ choose a refinement $\mathfrak{B}_1$ of $\mathfrak{B}$ of dimension $\leq n$ and let $\mathfrak{A}$ be a similar refinement of $\bar{\mathfrak{A}}^{2n+2}(\mathfrak{B}_1, \mathfrak{E}_1)$. Then the extension functions $\mathfrak{H}(\mathfrak{E}) = \bar{\mathfrak{H}}^{2n+2}(\mathfrak{E}_1)$ and $\mathfrak{A}(\mathfrak{B}, \mathfrak{E}) = \mathfrak{A}$ satisfy the definition of Clc.

For let K be any complex and τ' a partial realization of K on A of norm $< \mathfrak{H}$. We may assume that τ' is also a partial realization of K^{2n+2}, the $(2n + 2)$-section of K. Hence τ' can be extended to a realization τ_1 of K^{2n+2} on $A \cup B_1$, of norm $< \mathfrak{E}_1$. It follows from the fact that both A and B_1 are of dimension n at most, that $A \cup B_1$ cannot be more than $2n+1$ dimensional. As a consequence of this, τ_1 maps any $(2n+2)$-chain on K^{2n+2} into the null chain. Hence τ_1 can be extended to a mapping τ_2 defined on K by defining $\tau_2(Y^{2n+2+k})$ to be the null chain. Thus τ_2 is a realization of K on $A \cup B_1$ and τ_2 is of norm $< \mathfrak{E}$. Now let π be the projection of $A \cup B_1$ into $A \cup B$ which is obtained by projecting each vertex of B_1 into the corresponding vertex of B, leaving the vertices of A fixed. Then $\pi\tau_2 = \tau$ is a realization of K on $A \cup B$ of norm $< \mathfrak{E}$.

---

[8] Lebesgue dimension.

The above two theorems, together with the fact that in a compactum the Lebesgue definition of dimension is equivalent to the Menger-Urysohn definition, give a proof of Theorem 1. 2.

Because of the theorems of this section we can now drop the notation of q-Clc and use q-lc to mean that any one of the definitions 4. 1, 4. 2, or 4. 3 is satisfied.

Of the results in Section 2, Theorem 2. 1 generalizes easily to compact spaces. However Lemma 2. 4 depends essentially on the separability of $\mathfrak{S}$ and does not extend. Hence we have no proof that if $\mathfrak{S}$ is locally connected in dimensions $\leq n$ with respect to Čech cycles then it is lc^n. The converse is trivially true. For the case where the coefficient ring is a field, this equivalence has been proved by Čech [8], and this is sufficient for our later applications to generalized manifolds. The non-field case, however, remains open.

THEOREM 5. 5. *If $\mathfrak{S}$ is* lc^n, *then its homology groups $H^p(\mathfrak{S}, G)$, $p \leq n$, can be mapped isomorphically into the corresponding groups of a complex.*

Proof. Let $\mathfrak{E}$ be a covering and let $\mathfrak{U}_0$ be $* < \bar{\mathfrak{H}}^n(\mathfrak{E})$. To each Čech cycle, $\Gamma^p = \{X_\zeta^p\}$, let correspond the cycle $X_0^p = \phi(\Gamma^p)$ which is the coördinate on U_0 of Γ^p. The correspondence ϕ clearly induces a homomorphism Φ of $H^p(\mathfrak{S}, G)$ into $H^p(U_0, G)$.

Now suppose $\phi(\Gamma^p) = X_0^p \sim 0$ on U_0. Any $\mathfrak{U}_\zeta < \mathfrak{U}_0$, has a refinement $\mathfrak{U}_\rho < \mathfrak{A}^n(\mathfrak{U}_\zeta, \mathfrak{E})$. Let δ be the deformation chain associated with $\pi_\rho{}^0$ so that $F\delta X_\rho{}^p = X_\rho{}^p - \pi_\rho{}^0 X_\rho{}^p$. Let $Y_0{}^{p+1}$ and $Y'_0{}^{p+1}$ be chains of U_0 such that $FY_0{}^{p+1} = X_0{}^p$ and $FY'_0{}^{p+1} = \pi_\rho{}^0 X_\rho{}^p - X_0{}^p$. Then $Y^{p+1} = \delta X_\rho{}^p + Y_0{}^{p+1} + Y'_0{}^{p+1}$ is a chain of $U_\rho \cup U_0$ such that $FY^{p+1} = X_\rho{}^p$.

Let $Z^0{}_{0,\iota}$ be a vertex of $|Y^{p+1}|$ on U_0 and let $Z^0{}_{\rho,\iota}$ be a vertex of U_ρ contained in it. Define a partial realization of $|Y^{p+1}|$ on U_ρ by: $\tau' Y^q = Y^q$ if Y^q is on $|X_\rho{}^p|$, $\tau' Z^0{}_{0,\iota} = Z^0{}_{\rho,\iota}$. It is clear that norm $\tau' < St(\mathfrak{U}_0; \mathfrak{U}_0) < \bar{\mathfrak{H}}^n(\mathfrak{E})$. By the choice of $\mathfrak{U}_\rho$, τ' can then be extended to a realization τ of $|Y^{p+1}|$ on $U_\rho \cup U_\zeta$ and τY^{p+1} is such that $F\tau Y^{p+1} = X_\rho{}^p$. Now let π be the obvious projection of $U_\rho \cup U_\zeta$ onto U_ζ. Then $\pi\tau Y^{p+1}$ is a chain on U_ζ and $F\pi\tau Y^{p+1} = \pi_\rho{}^\zeta X^p \sim X_\zeta{}^p \sim 0$. Hence $\Gamma^p \sim 0$ which means that Φ is an isomorphism.

6. Local Betti numbers. In the definition of generalized manifolds two local invariants are used, local connectedness and local Betti numbers. In this section we study the latter and show that they can be expressed in terms of cocycles. For this it is necessary that the coefficient ring be a field, F, which we now assume.

For the definition of cocycles of $\mathfrak{S}$, we need the fact[9] that if σ is a simplicial mapping, then the inverse of σ, which we denote by σ^*, defines a cochain-mapping which commutes with the boundary operator. Using this, a Čech cocycle[10] Γ_q is defined as a collection $\{X_q^\zeta\}$ of cocycles on the nerves of some family $\{\mathfrak{U}_\zeta\}$ of coverings of $\mathfrak{S}$ such that if $\mathfrak{U}_\rho < \mathfrak{U}_\zeta$ and if Γ_q has coördinates on both U_ζ, U_ρ, then $\pi_\rho^{\zeta *} X_q^\zeta \sim X_q^\rho$ on U_ρ. By definition, $\Gamma_q \sim 0$ if on some U_ρ, $\pi_\rho^{\zeta *} X_q^\zeta \sim 0$. The cohomology groups $H_q(\mathfrak{S}, G)$ of $\mathfrak{S}$ are defined in the usual way.

Let Γ_q be defined on U_ζ, but not on U_ρ. Then if we add to the coördinates of Γ_q the coördinate $\pi_\rho^{\zeta *} X_q^\zeta$, this new collection of cocycles is a Čech cocycle, and is homologous to Γ_q. Hence we can always assume that if Γ_q has a coördinate on U_ζ, it also has coördinates on each U_ρ, for $\mathfrak{U}_\rho < \mathfrak{U}_\zeta$.

We say that a cochain on a nerve A is in an open set if the set contains the kernel of each simplex on which the cochain is defined. A cocycle $\Gamma_q = \{X_q^\zeta\}$ is said to be in an open set if each X_q^ζ is in the set and if also each of the cochains involved in the cohomologies $\pi_\rho^{\zeta *} X_q^\zeta \sim X_q^\rho$ is in the set. Note that if $\mathfrak{U}_\rho < \mathfrak{U}_\zeta$ and if X_q^ζ is in an open set, then $\pi_\rho^{\zeta *} X_q^\zeta$ is in the same set. Hence if Γ_q is in an open set and if we add new coördinates to Γ_q as above, the new cocycle is in the same set.[11]

Local Betti numbers have been studied by Alexandroff [1], Čech [7], and others. For our purposes, Čech's definition is the most convenient. Let $P \supset Q$ be open sets of S which contain the point s. Denote by $H^q(P, Q, s)$ the group of q-cycles mod $\mathfrak{S} - P$ reduced by those which are ~ 0 mod $\mathfrak{S} - Q$, and let $R^q(P, Q, s)$ be the rank of this group. For $Q' \subset Q$, $R^q(P, Q, s) \geq R^q(P, Q', s)$, so there is a greatest cardinal number $R^q(P, s) \leq R^q(P, Q, s)$ for all neighborhoods Q of s in P. For $P \supset P' \supset Q$, $R^q(P, Q, s) \leq R^q(P', Q, s)$, so there is a least cardinal number $R^q(s) \geq R^q(P, s)$ for all neighborhoods P of s. The number $R^q(s)$ is called the local Betti number of $\mathfrak{S}$ at the point s.

---

[9] See, for example, [19, p. 401].

[10] We follow Lefschetz [15] in considering cocycles to be cycles in the dual complex. The chains on the dual are called cochains, and their dimensional indices are written as subscripts rather than superscripts. F is used for the boundary operator on both the complex and its dual.

[11] For a further discussion of Čech cycles see [10].

Now let the group $H^q(P, Q, s)$ be replaced by $H_q(P, Q, s)$, the group of q-cocycles in Q reduced by those which are cohomologous to zero in P. Let $R_q(P, Q, s)$ be the rank of $H_q(P, Q, s)$. Then it is easily seen that $R_q(P, s) \leq R_q(P, Q, s)$ exists as above, and also $R_q(s) \geq R_q(P, s)$. The number $R_q(s)$ is called the local co-Betti number of $\mathfrak{S}$ at s.

A pair of neighborhoods, P and Q, of s such that $R_q(P, Q, s) = R_q(s)$ is called a *canonical* pair. It is clear that any neighborhood of s contains a canonical pair if $R_q(s)$ is finite. Hence if $\mathfrak{C}$ is any covering of $\mathfrak{S}$, and if $R_q(s)$ has the same finite value, R_q, for each point s of $\mathfrak{S}$, then there are refinements $\mathfrak{G}_q = \mathfrak{G}_q(\mathfrak{C})$ and $\mathfrak{H}_q = \mathfrak{H}_q(\mathfrak{C})$ where each element of $\mathfrak{H}_q$ is a Q of a canonical pair and each element of $\mathfrak{G}_q$ is a P corresponding to one of the Q's of $\mathfrak{H}_q$.

Definition 6.1. $\mathfrak{S}$ is q-co-lc if for each covering $\mathfrak{C}$ there is a refinement $\mathfrak{H}_q = \mathfrak{H}_q(\mathfrak{C})$ such that if Γ_q is a cocycle in a set of $\mathfrak{H}_q$, then $\Gamma_q \sim 0$ in a set of $\mathfrak{C}$. If $\mathfrak{S}$ is q-co-lc for all $q \leq n$, we say that $\mathfrak{S}$ is lc$_n$.

This definition is clearly equivalent to

Definition 6.2. $\mathfrak{S}$ is q-co-lc if $R_q(s) = 0$ for every $s \in \mathfrak{S}$.

In order to compare local Betti numbers and local co-Betti numbers we make use of an intersection between cycles and cocycles defined[12] by Lefschetz [15]. This is an extension of the intersections defined on complexes by Čech [9] and Whitney [19]. If Γ^p and Γ_q, $q \leq p$, are a cycle and a cocycle, the intersection, denoted by $(\Gamma^p \cdot \Gamma_q)$ is a cycle of dimension $p - q$.

If Γ_q is defined on U_ζ, then the coördinate on this nerve of $(\Gamma^p \cdot \Gamma_q)$ is the intersection (= Whitney's cap product) of Γ_ζ^p and Γ_q^ζ. This is a cycle of U_ζ. Since we may assume that Γ_q has coördinates on a cofinal family of coverings of $\mathfrak{S}$, we have the intersection defined on this cofinal family and hence, in the usual way, on every covering. That this collection of cycles is actually a cycle of $\mathfrak{S}$ follows immediately from theorem 6 of [19].

Lefschetz has proved the following:[13]

---

[12] The definition makes use of a generalization of (5.14) of [19].

[13] For field coefficients, this theorem can be proved as follows: Suppose that there is a $\Gamma^p \bmod \mathfrak{S} - Q \not\sim 0 \bmod \mathfrak{S} - Q$. Then on some U_ζ, $\Gamma_\zeta^p \not\sim 0 \bmod \mathfrak{S} - Q$. Hence, by the theorem for complexes, there is a cocycle Γ_p^ζ on U_ζ in Q such that $KI(\Gamma_\zeta^p \cdot \Gamma_p^\zeta) = 1$. But Γ_p^ζ is also a Čech cocycle, Γ_p, in Q and $KI(\Gamma^p \cdot \Gamma_p) \neq 0$.

Now suppose that there is a cocycle Γ_p in $Q \not\sim 0$ in Q. Then there is a $\mathfrak{U}_\zeta$ such that $\Gamma_p^\zeta \not\sim 0$ in U_ζ. Let $\mathfrak{U}_\rho$ be a normal refinement [4] of $\mathfrak{U}_\zeta$. On U_ρ we have $\pi_\rho^{\zeta *}\Gamma_p^\zeta \not\sim 0$ in Q. Hence, by the theorem for complexes, there is a cycle $\Gamma_\rho^p \bmod \mathfrak{S} - Q$ such that $KI(\Gamma_\rho^p \cdot \pi_\rho^{\zeta *} \Gamma_p^\zeta) = 1$. Then $KI(\pi_\rho^\zeta \Gamma_\rho^p \cdot \Gamma_p^\zeta) = 1$. Since $\mathfrak{U}_\rho$ is a normal refinement of $\mathfrak{U}_\zeta$, $\pi_\rho^\zeta \Gamma_\rho^p$ is a coordinate of a Čech cycle, Γ^p, of $\mathfrak{S} \bmod \mathfrak{S} - Q$, and $KI(\Gamma^p \cdot \Gamma_p) \neq 0$.

THEOREM. *For any open set Q in $\mathfrak{S}$, a necessary and sufficient condition that a cycle Γ^p mod $\mathfrak{S} - Q$ be ~ 0 mod $\mathfrak{S} - Q$, is that for every cocycle Γ_p in Q, $KI(\Gamma^p \cdot \Gamma_p) = 0$, and a necessary and sufficient condition that a cocycle Γ_p in Q be ~ 0 in Q is that for every cycle Γ^p mod $\mathfrak{S} - Q$, $KI(\Gamma^p \cdot \Gamma_p) = 0$.*

An immediate consequence of this theorem is that $H^p(\mathfrak{S}, F) = H_p(\mathfrak{S}, F)$. Also, by means of this theorem we prove

THEOREM 6.1. *For any pair P and Q of neighborhoods of s, $R^q(P, Q, s) = R_q(P, Q, s)$.*

Proof. Denote these numbers by R^q and R_q. Let $\Gamma_1{}^q, \cdots, \Gamma_t{}^q$ be a set of cycles mod $\mathfrak{S} - P$ which are independent mod $\mathfrak{S} - Q$. Considered as cycles mod $\mathfrak{S} - Q$, they are still independent. By Lefschetz's theorem, there exist cocycles $\Gamma_q{}^1, \cdots, \Gamma_q{}^t$ in Q such that the determinant $| a_i{}^j |$ is not zero, where $a_i{}^j = KI(\Gamma_i{}^q \cdot \Gamma_q{}^j)$.

Considering these cocycles as being in P, the intersection numbers $KI(\Gamma_i{}^q \cdot \Gamma_q{}^j)$ remain the same. If $\Gamma_q{}^1, \cdots, \Gamma_q{}^t$ were not independent in P, then the determinant $| a_i{}^j |$ would be zero, which is impossible. Hence there are at least t cocycles in Q which are independent in P, i. e., $R^q \leq R_q$.

But an interchange of P and Q and of cycles and cocycles in the above argument shows that $R^q \geq R_q$; i. e. $R^q = R_q$.

An immediate consequence of this is

THEOREM 6.2. *For each point s of $\mathfrak{S}$, $R^q(s) = R_q(s)$.*

7. Generalized manifolds. The definition of a generalized manifold which we now introduce is similar to Čech's [5, 7], except that we use cocycles and local co-Betti numbers instead of relative cycles and local Betti numbers. It is this that makes possible the rapid proof of the Poincaré Duality Theorem.

Definition 7.1. A compact space $\mathfrak{M}$ is said to be a generalized n-manifold over a coefficient field F if

a) dimension $\mathfrak{M} = n$; b) $\mathfrak{M}$ is lc^n; c) $\mathfrak{M}$ is lc$_{n-1}$; d) $R_n(s) = 1$ for each point s of $\mathfrak{M}$.

Since $\mathfrak{M}$ is compact and lc^0, it has at most a finite number of components, and each of these has properties a) $\cdots$ d). Hence we can assume without any loss of generality that $\mathfrak{M}$ is connected.

The easiest way to define orientability of a manifold is to require that there be an n-cycle on the manifold $\not\sim 0$ and that on any proper closed subset all n-cycles are ~ 0. However, for any applications it is better to have a definition with a more local character. This is done as follows:

Let $\mathfrak{E}$ be any covering of $\mathfrak{M}$ and let $\mathfrak{G}_n(\mathfrak{E})$, $\mathfrak{H}_n(\mathfrak{E})$ be the refinements of $\mathfrak{E}$ introduced in Section 6. In each set Q_i of $\mathfrak{H}^n$ there is an n-cocycle $\Gamma_n{}^i \not\sim 0$ in P_i of G_n. Suppose Q_i and Q_j have a non-vacuous intersection. At any point of this intersection a canonical pair $P \supset Q$ can be chosen such that P is in the intersection and in Q there is an n-cocycle $\Gamma_n \not\sim 0$ in P. Since Γ_n is then in Q_i, $\Gamma_n{}^i$ and Γ_n are dependent in P_i, so we have $f_i\Gamma_n{}^i \sim \Gamma_n$ in P_i. Similarly $f_j\Gamma_n{}^j \sim \Gamma_n$ in P_j.

We say that $\mathfrak{M}$ is *orientable* if for each $\mathfrak{E}$ and for each pair of refinements $\mathfrak{G}_n$, $\mathfrak{H}_n$, the cocycles $\Gamma_n{}^i$ can be so chosen that for any choice of canonical pairs in the intersections of the sets of $\mathfrak{H}_n$, the ratios f_i/f_j are all $+1$. The $\Gamma_n{}^i$ so chosen are called fundamental cocycles.

THEOREM. 7. 1. *If $\mathfrak{M}$ is an orientable generalized n-manifold, there is an n-cycle Γ^n on M which is not homologous to any cycle on any proper closed subset of $\mathfrak{M}$. Conversely, if such a cycle exists, $\mathfrak{M}$ is orientable.*

Proof. Let $\mathfrak{E}$ be any covering of $\mathfrak{M}$ and let $\mathfrak{H}^{(1)} = \mathfrak{H}_n(\mathfrak{E})$. By 4(b) there exists a covering $\mathfrak{H}^{(2)} \ll \mathfrak{H}^{(1)}$. Let $\mathfrak{U}$ be such that $St(\mathfrak{H}^{(2)}; \mathfrak{U}) < \mathfrak{H}^{(1)}$, and let $\{\mathfrak{U}_\zeta\}$ be a complete family of coverings which are refinements of $\mathfrak{U}$ and are n-dimensional. All cycles and cocycles are to be considered only on this family.

Now let $Q_i{}^1$, $Q_j{}^1$ be two sets of $\mathfrak{H}^{(1)}$ and let $\Gamma_n{}^i$, $\Gamma_n{}^j$ be the fundamental cocycles in these sets. Let $Q_i{}^2$, $Q_j{}^2$ be the corresponding sets of $\mathfrak{H}^{(2)}$ and suppose that $\bar{Q}_i{}^2 \cap \bar{Q}_j{}^2 \neq 0$. At any point of this intersection there is a canonical pair, $P \supset Q$, such that P is in $Q_i{}^1 \cap Q_j{}^1$, and in any such Q there is a cocycle Γ_n. Since $\mathfrak{M}$ is orientable, Γ_n may be so chosen that $\Gamma_n \sim \Gamma_n{}^i$ in E_i, $\Gamma_n \sim \Gamma_n{}^j$ in E_j, where E_i, E_j are sets of $\mathfrak{E}$ containing $Q_i{}^1$, $Q_j{}^1$.

By Theorem 6. 1, there is a cycle $\Gamma_i{}^n$ mod $\mathfrak{M} - E_i$ such that $KI(\Gamma_i{}^n \cdot \Gamma_n{}^i) = 1$. $\Gamma_i{}^n$ is also a cycle mod $\mathfrak{M} - P$, and since $\Gamma_n{}^i \sim \Gamma_n$, we have $KI(\Gamma_i{}^n \cdot \Gamma_n) = 1$. Similarly, there is a cycle $\Gamma_j{}^n$ mod $\mathfrak{M} - E_j$ such that $KI(\Gamma_j{}^n \cdot \Gamma_n) = 1$. Since P and Q are a canonical pair, it follows that $\Gamma_i{}^n \sim \Gamma_j{}^n$ mod $\mathfrak{M} - P$. Furthermore, each U_ζ being n-dimensional, homology for n-cycles means identity, so we have $\Gamma_i{}^n = \Gamma_j{}^n$ mod $\mathfrak{M} - P$. This is true for each point of $\bar{Q}_i{}^2 \cap \bar{Q}_j{}^2$, so that we have $\Gamma_i{}^n = \Gamma_j{}^n$ on this set.

Now we can proceed to the construction of Γ^n. For this we need coverings $\mathfrak{H}^{(4)} \ll \mathfrak{H}^{(3)} \ll \mathfrak{H}^{(2)}$. We assume that $\mathfrak{U}$ is $< \mathfrak{H}^{(4)}$ and such that $St(\mathfrak{H}^{(4)}; \mathfrak{U}) < \mathfrak{H}^{(3)}$ and $St(\mathfrak{H}^{(3)}; \mathfrak{U}) < \mathfrak{H}^2$. We first define on each U_ζ an n-cycle $X_\zeta{}^n$ as follows: Let $Z^n{}_{\zeta,k}$ be any n-simplex of U_ζ. It is on some $\bar{Q}_i{}^3$. In $X_\zeta{}^n$ let it have the coefficient which it has in $X^n{}_{\zeta,i}$ where $X^n{}_{\zeta,i}$ is the coördinate of $\Gamma_i{}^n$ on U_ζ. That this assignment of coefficients is without ambiguity results from the fact that if $Z^n{}_{\zeta,k}$ is on $\bar{Q}_\iota{}^3$ and on $\bar{Q}_j{}^3$, then it is also on $\bar{Q}_\iota{}^2 \cap \bar{Q}_j{}^2$, because of the choice of $\mathfrak{U}$ and of $\mathfrak{H}^{(3)}$, and on this set $\Gamma_\iota{}^n = \Gamma_j{}^n$.

To see that $X_\zeta{}^n$ is a cycle, let $Z^{n-1}{}_{\zeta,j}$ be any $(n-1)$-simplex of U_ζ. It is on some $\bar{Q}_\iota{}^4$ and every n-simplex of which it is a face is on $\bar{Q}_\iota{}^3$ by the choice of $\mathfrak{H}^{(4)}$ and $\mathfrak{U}$. These simplexes have the same coefficients in $X_\zeta{}^n$ as they have in $X^n{}_{\zeta,\iota}$, and $FX^n{}_{\zeta,\iota}$ is on $\mathfrak{M} - E_\iota$. $Z^{n-1}{}_{\zeta,j}$ has therefore a zero coefficient in $FX_\zeta{}^n$, so $X_\zeta{}^n$ is a cycle.

Let $\Gamma^n = \{X_\zeta{}^n\}$. To see that Γ^n is a Čech cycle, we need merely notice that if $\mathfrak{U}_\rho < \mathfrak{U}_\zeta$, then $\pi_\rho{}^\zeta X^n{}_{\rho,\iota} = X^n{}_{\zeta,\iota}$, so $\pi_\rho{}^\zeta X_\rho{}^n = X_\zeta{}^n$.

Finally let M be any proper closed subset of $\mathfrak{M}$. In $\mathfrak{M} - M$ there is a canonical pair $P \supset Q$, and in Q an n-cocycle, $\Gamma_n \not\sim 0$ in P. If Γ^n is on M, then $KI(\Gamma^n \cdot \Gamma_n) = 0$. But by construction, $KI(\Gamma^n \cdot \Gamma_n) \neq 0$, and so Γ^n cannot be on M.

For the converse, all that is necessary is to choose the $\Gamma_n{}^\iota$ so that $KI(\Gamma^n \cdot \Gamma_n{}^\iota) = +1$ for all i.

Theorem 7.2. *If $\mathfrak{M}$ is not orientable, then every n-cycle on $\mathfrak{M}$ is ~ 0.*

Proof. It is sufficient to show that every n-cocycle on $\mathfrak{M}$ is ~ 0. Now if $\mathfrak{M}$ is not orientable, there is a covering $\mathfrak{E}$ and one of its refinements $\mathfrak{H}_n(\mathfrak{E})$ such that there is a sequence $\Gamma_n{}^1 \sim \Gamma_n{}^2 \sim \cdot \cdot \cdot \sim \Gamma_n{}^k \sim f\Gamma_n{}^1$, $f \neq 1$, where the $\Gamma_n{}^\iota$ are given by 7(d). Then, since the coefficient group is a field $\Gamma_n{}^1 \sim 0$ in $\mathfrak{M}$ and hence, since $\mathfrak{M}$ is connected, every $\Gamma_n{}^\iota \sim 0$ in $\mathfrak{M}$.

Now let Γ_n be an arbitrary n-cocycle on $\mathfrak{M}$. Γ_n has a coördinate $X_n{}^\zeta$ on the nerve of an n-dimensional $\mathfrak{U}_\zeta$ which is $* < \mathfrak{H}_n(\mathfrak{E})$. Since U_ζ is n-dimensional, each n-simplex is an n-cocycle, $W_n{}^{\zeta,\iota}$, and $X_n{}^\zeta = \Sigma f_\iota W_n{}^{\zeta,\iota}$. But also each $W_n{}^{\zeta,\iota}$ is a Čech cocycle, $W_n{}^{\zeta,\iota}$, and we have $\Gamma_n \sim \Sigma f_\iota W_n{}^\iota$. Since each $W_n{}^\iota$ is in a set of $\mathfrak{H}_n(\mathfrak{E})$, each $W_n{}^\iota$ is homologous to a $\Gamma_n{}^\iota$ and hence is ~ 0 in $\mathfrak{M}$. But then $\Gamma_n \sim 0$ in $\mathfrak{M}$ also, which completes the proof.

Theorem 7.3. *If $\mathfrak{M}$ is orientable, the Poincaré duality theorem holds in* $\mathfrak{M}$, i. e., $H^p(\mathfrak{M}, F) = H^{n-p}(\mathfrak{M}, F)$.

Proof. Let Γ_p be any cocycle on $\mathfrak{M}$. Then $(\Gamma^n \cdot \Gamma_p) = \Gamma^{n-p}$ is a cycle on $\mathfrak{M}$. Set $\phi(\Gamma_p) = \Gamma^{n-p}$. Since $\phi(\Gamma_p) \sim 0$ if $\Gamma_p \sim 0$, ϕ induces a homomorphism Φ of $H_p(\mathfrak{M}, F)$ into $H^{n-p}(\mathfrak{M}, F)$. Since $H_p(\mathfrak{M}, F) = H^p(\mathfrak{M}, F)$, there is a homomorphism Φ_1 of $H^p(\mathfrak{M}, F)$ into $H^{n-p}(\mathfrak{M}, F)$. It is shown below that this is a homomorphism "onto." But this means that rank $H^p(\mathfrak{M}, F)$ $\geq$ rank $H^{n-p}(\mathfrak{M}, F)$. If p is replaced by $n-p$, the inequality runs in the other direction, so that these two groups have the same rank. Hence they are isomorphic.

To show that Φ maps $H_p(\mathfrak{M}, F)$ onto $H^{n-p}(\mathfrak{M}, F)$, it is necessary to show that for any Γ^{n-p} there is a Γ_p such that $(\Gamma^n \cdot \Gamma_p) \sim \Gamma^{n-p}$. Before proceeding to the construction of such cocycles we first note that by an argument analogous to that used in proving Theorem 1. 1 we can show that any covering $\mathfrak{E}$ has a refinement $\underline{\mathfrak{H}}_n(\mathfrak{E})$ for which the following statement is true: Let τ'^* assign to each vertex of a complex K of dimension $\leq n$ a coördinate, on a fixed nerve U_2, of one of the fundamental cocycles of $\mathfrak{M}$ in such a way that if $Z_i^0, \cdots, Z_j^0$ are the vertices of a simplex of K, then $\tau'^*Z_i^0 + \cdots + \tau'^*Z_j^0$ is in a single set of $\underline{\mathfrak{H}}_n(\mathfrak{E})$. Then there is a τ^* which assigns to each q-simplex of K an $(n - q)$-cochain on a U_3, where $\mathfrak{U}_3 < \mathfrak{U}_2$, in such a way that 1) $\tau^*FZ_i^q = F\tau^*Z_i^q$, 2) for each simplex Z_i^q of K there is a set of $\mathfrak{E}$ which contains $\tau^*Z_i^q$, and 3) for each vertex of K, $\tau^*Z_i^0 = \pi_3^{2*}\tau'^*Z_i^0$.

Now let $\mathfrak{E}$ be any covering of $\mathfrak{M}$ and choose $\mathfrak{E}_1^* < \bar{\mathfrak{H}}^n(\mathfrak{E})$. Let $\mathfrak{U}_0$ be the refinement of $\mathfrak{E}$ introduced in the proof of theorem 5. 5, and let $\mathfrak{U}_1$ be an n-dimensional refinement of $\mathfrak{U}_0$ which is $^* < \underline{\mathfrak{H}}_n(\mathfrak{E})$ and $< \bar{\mathfrak{A}}^n(\mathfrak{U}_0, \mathfrak{E})$. In each set of $\mathfrak{U}_1$ there is a fundamental n-cocycle of $\mathfrak{M}$. There being only a finite number of vertices of U_1, we may assume that all of these cocycles have coördinates on a U_2, where $\mathfrak{U}_2 < \mathfrak{U}_1$. Let τ'^* assign to each vertex of U_1 the corresponding cocycle on U_2. Then τ'^* satisfies the conditions above, so that there exists a τ^* with the above properties.

Let $\Gamma^{n-p} = \{X_\zeta^{n-p}\}$ be any Čech cycle on $\mathfrak{M}$. Then $\tau^*X_1^{n-p} = X_p^3$ is a cocycle on U_3. This generates a Čech p-cocycle $\Gamma_p = \{X_p^\zeta\}$, where $X_p^\zeta = \pi_\zeta^{3*}X_p^3$ for any $\mathfrak{U}_\zeta < \mathfrak{U}_3$. Let $(\Gamma^n \cdot \Gamma_p) = \{W_\zeta^{n-p}\}$.

To show that $(\Gamma^n \cdot \Gamma_p) \sim \Gamma^{n-p}$, it is sufficient by Theorem 5. 5 to show that the coördinates on U_0 of these two cycles are homologous on U_0. Let L be the cartesian product of U_1 with a 1-simplex, subdivided simplicially in such a way that all the vertices of L are among those of the base and the top. We identify the base of L with U_1. We note that if X^q is a cycle on U_1 and if $\bar{X}^q$ is the corresponding cycle on the top of L, then there is a chain Y^{q+1} in L such that $FY^{q+1} = X^q - \bar{X}^q$.

We define a mapping τ' as follows: On U_1, τ' is the identity and for any chain $\bar{Y}^q$ on the top of L, $\tau'\bar{Y}^q = \pi^1_3(X_3^n \cdot \tau^*Y^q) = \pi^1_3W_3^q$, where X_3^n is the coördinate of Γ^n on U_3. τ' is clearly a partial realization of L on U_2, and norm $\tau' < \bar{\mathfrak{H}}^n(\mathfrak{E})$. Hence τ' can be extended to a realization τ of L on $U_1 \cup U_0$.

Let Y^{n-p+1} be the chain on L such that $FY^{n-p+1} = X_1^{n-p} - \bar{X}_2^{n-p}$, where X_1^{n-p} is the coördinate of Γ^{n-p} on U_1. Then $F\tau Y^{n-p+1} = X_1^{n-p} - \pi^1_3W_3^{n-p}$, and τY^{n-p+1} is on $U_1 \cup U_0$. If π is the obvious projection of $U_1 \cup U_0$ onto U_0, then $\pi\tau Y^{n-p+1}$ is on U_0 and $F\pi\tau Y^{n-p+1} = \pi_1^0X_1^{n-p} - \pi_1^0\pi_3^1W_3^{n-p}$. But this suffices to show that $X_0^{n-p} \sim W_0^{n-p}$ on U_0, and hence that $\Gamma^{n-p} \sim (\Gamma^n \cdot \Gamma_p)$, which completes the proof.

8. Čech manifolds. A compact[14] space $\mathfrak{C}$ is a Čech n-manifold if the following axioms are satisfied:

8. a) $\dim \mathfrak{C} = n$.

8. b) $\mathfrak{C}$ is lc^n.

8. c) $R^q(s) = 0$ for all $q < n$ and for all $s \in \mathfrak{C}$.

8. d) $R^n(s) = 1$ for each point s of $\mathfrak{C}$.

8. e) If M is any proper closed subset of $\mathfrak{C}$ and if s is a point on the boundary of M, then $R_M^n(s) = 0$, where $R_M^n(s)$ is the local Betti number of the space M at the point s.

$\mathfrak{C}$ is said to be *orientable* if there exists an n-cycle on $\mathfrak{C}$ which is not homologous to any cycle on any proper closed subset of $\mathfrak{C}$.

THEOREM 8. 1. *A Čech n-manifold is a generalized n-manifold and conversely.*

Proof. By Theorem 6. 2, 8. c) and 8. d) are equivalent to 7. c) and 7. d). Hence we need only show that 8. e) holds in a generalized n-manifold. By Theorem 6. 2, $R_M^n(s) = R_n^M(s)$, where the latter number is the local co-Betti number of the space M at the point s. It is sufficient therefore to show that any neighborhood N in M of a point s on the boundary of M contains a neighborhood such that any n-cocycle Γ_n in the smaller is homologous to zero in the larger. N is the intersection with M of an open set in $\mathfrak{C}$, and in this open set there exists a canonical pair $P \supset Q \supset s$. Let P_1 and Q_1 be the intersections of P and Q with M, and $P_2 = P - P_1$, $Q_2 = Q - Q_1$. We show that any n-cocycle of M in Q_1 is ~ 0 in N. Since M is closed, P_2 and Q_2 are open. Hence in Q_2 there is a canonical pair containing an n-cocycle $\Gamma_{n,2}$ such that $KI(\Gamma^n \cdot \Gamma_{n,2}) = 1$.

Let $\Gamma_n = \{X_n^\zeta\}$ be any cocycle of the space M in Q_1. We note that since M is not necessarily open, Γ_n may not be a cocycle of the space $\mathfrak{C}$. Let $\mathfrak{U}_\zeta$ be an n-dimensional covering such that both Γ_n and $\Gamma_{n,2}$ have coördinates on U_ζ. Since U_ζ is n-dimensional, X_n^ζ is a cocycle not only of $M \cap U_\zeta$, but also of U_ζ. Since X_n^ζ and $X^\zeta_{n,2}$ are both in Q, and $\Gamma_{n,2} \not\sim 0$, we have $X_n - fX_{n,2} \sim 0$ in P.

---

[14] Čech also assumes that the space is perfectly normal, but this is not necessary for the compact space.

Hence there is a $\mathfrak{U}_\rho < \mathfrak{U}_\zeta$ such that $\pi_\rho{}^{\zeta*}(X_n{}^\zeta - fX^\zeta{}_{n,2}) = FY^\rho{}_{n-1}$, where $Y^\rho{}_{n-1}$ is in P. Let $Y^\rho{}_{n-1,2}$ be the part of $Y^\rho{}_{n-1}$ in P_2, and $Y^\rho{}_{n-1,1} = Y^\rho{}_{n-1} - Y^\rho{}_{n-1,2}$. Since P_2 is open, $FY^\rho{}_{n-1,2}$ is in P_2 and so is $\pi_\rho{}^{\zeta*}X^\zeta{}_{n,2}$. Let $X^\rho{}_{n,1} = FY^\rho{}_{n-1,2} + \pi_\rho{}^{\zeta*}fX^\zeta{}_{n,2}$. Then $FY^\rho{}_{n-1,2} = \pi_\rho{}^{\zeta*}X_n{}^\zeta - X^\rho{}_{n,1}$. But then the part of $FY^\rho{}_{n-1,1}$ in P_1 is just $\pi_\rho{}^{\zeta*}X_n{}^\zeta$. Hence $\Gamma_n \sim 0$ in $P_1 \subset N$, which completes the proof.

9. Examples. There is a well-known example to show that homology local connectedness is weaker than homotopy local connectedness. This consists of a sequence of Poincaré spaces [16, p. 218] converging to a point and joined by simple arcs whose diameters converge to zero. This is locally connected in dimension 0 in any sense. Since the fundamental group of a Poincaré space does not vanish, there are arbitrarily small continuous 1-spheres in this space which cannot be shrunk to points, and hence the space is not locally connected in dimension 1 in the homotopy sense. However, the 1-dimensional homology group of a Poincaré space consists of merely the null element, so this space is lc^1.

To see the effect of different coefficient rings on local connectedness consider a sequence of projective planes which converge to a point and are connected by simple arcs as above. A projective plane is a 2-cycle mod 2, and so this space is not 2-lc for mod 2 coefficients. However, there are no integral 2-cycles on the space, and so for integer coefficients it is 2-lc. Each of these projective planes has a 1-dimensional torsion cycle, and so for both integer and mod 2 coefficients the space is not 1-lc, while for rational coefficients it is 1-lc.

The following example, due to van Kampen, shows that a generalized manifold need not be either a topological or a combinatorial manifold. Let E be Euclidean 3-space, made compact by adding the point at infinity. From E remove a sequence $\{A_i\}$ of open 3-cells which converge to a point. Then from each of a sequence $\{P_i\}$ of Poincaré spaces remove an open 3-cell, B_i. Now identify the boundary in E of the cell A_i with the boundary in P_i of the cell B_i. The resulting space, which is a compactum, is a generalized manifold. However, it is not locally connected in dimension 1 in the homotopy sense, and so it is neither a topological nor a combinatorial manifold.

PRINCETON UNIVERSITY.

BIBLIOGRAPHY

1. P. Alexandroff, "On local properties of closed sets," *Annals of Mathematics*, vol. 36 (1935), pp. 1-35.

2. P. Alexandroff and H. Hopf, *Topologie*, vol. 1, Berlin (1935).

3. P. Alexandroff, "Zur Homologie-Theorie der Kompakten," *Compositio Mathematica*, vol. 4 (1936), pp. 256-270.

4. E. Čech, "Théorie générale de l'homologie dans un espace quelconque," *Fundamenta Mathematicae*, vol. 19 (1932), pp. 149-183.

5. E. Čech, "Théorie générale des variétés et de leurs théorèmes de dualité," *Annals of Mathematics*, vol, 34 (1933), pp. 621-730.

6. E. Čech, "Contribution à la théorie de la dimension," *Cas. Mat. Fys.*, vol. 62 (1933), pp. 277-291.

7. E. Čech, "Sur les nombres de Betti locaux," *Annals of Mathematics*, vol. 35 (1934), pp. 678-701.

8. E. Čech, "Sur la connexité locale d'ordre supérieur," *Compositio Mathematica*, vol. 2 (1935), pp. 1-25.

9. E. Čech, "Multiplications on a complex," *Annals of Mathematics*, vol. 37 (1936), pp. 681-697.

10. C. H. Dowker, "Mapping theorems in non-compact spaces," Thesis, Princeton (1938).

11. S. Lefschetz, "On generalized manifolds," *American Journal of Mathematics*, vol. 55 (1933), pp. 469-504.

12. S. Lefschetz, "On locally connected and related sets," *Duke Mathematical Journal*, vol. 2 (1936), pp. 435-442.

13. S. Lefschetz, "On the fixed point formula," *Annals of Mathematics*, vol. 38 (1937), pp. 819-822.

14. S. Lefschetz, "On chains of topological spaces," *Annals of Mathematics*, vol. 39 (1938), pp. 383-396.

15. S. Lefschetz, *Topology*, second edition. New York, 1942.

16. H. Seifert and W. Threlfall, "*Lehrbuch der Topologie*," Leipzig und Berlin (1934).

17. J. Tukey, "Convergence and uniformity in topology," *Annals of Mathematical Studies*, II, (1940).

18. L. Vietoris, "Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen," *Math. Ann.*, vol. 97 (1927), pp. 454-472.

19. H. Whitney, "On products in a complex," *Annals of Mathematics*, vol. 39 (1938), pp. 397-432.

20. R. L. Wilder, "Generalized closed manifolds in n-space," *Annals of Mathematics*, vol. 35 (1934), pp. 876-903.

21. R. L. Wilder, "On locally connected spaces," *Duke Mathematical Journal*, vol. 1 (1935), pp. 543-555.

ON CONVERGENCE AND SUMMABILITY OF TRIGONOMETRIC SERIES.*

By OTTO SZÁSZ.

1. It is well known that a Fourier cosine series

$$\phi(t) \sim \sum_{1}^{\infty} a_n \cos nt = \sum a_n \cos nt$$

can be integrated termwise, i. e.,

$$\Phi(t) = \int_0^t \phi(u)\,du = \sum n^{-1} a_n \sin nt,$$

and the series on the right is uniformly convergent; $\Phi(t)$ is absolutely continuous. This gives rise to the following definition of summability:

We say that a series $\sum a_n$ is summable to the value s by Lebesgue's method, if the series

(1.1) $$\sum n^{-1} a_n \sin nt = F(t)$$

converges in an interval $0 < t < t_0$, and if

(1.2) $$t^{-1} F(t) \to s \quad \text{as} \quad t \downarrow 0.$$ [Cf. 6, pp. 272, 303, 304].[1]

We then write $L \sum a_n = s$.

It is known that convergence does not imply L-summability, and, conversely, the latter does not imply convergence. Thus an additional condition is needed in order to deduce convergence from L-summability, and conversely. This condition may be imposed upon the terms a_n, or upon the function $F(t)$, or upon both; the problem is of the well known Tauberian type. Similar problems arise, if we replace convergence by bounds of oscillation, or just by boundedness.

If $\sum a_n \cos nt$ is a Fourier series, and the corresponding even function is denoted by $\phi(t)$, then $F(t) = \int_0^t \phi(u)\,du$; thus in this case L-summability of $\sum a_n$ means continuity in the mean of $\phi(t)$ at $t = 0$.

Note that the summability method L is a series to function transform with the convergence factors $\sin nt/nt$, $n = 1, 2, 3, \cdots$.

---

* Presented to the Society, April 12, 1941; Received June 1, 1941.

[1] Numbers in square brackets refer to the bibliography at the end of this paper.

We can increase the scope of L-summability by associating with the series (1.1) a function $F(t)$ in a different way, e.g. instead of convergence assume summability by Abel's method [Cf. Verblunsky 4], or assume that (1.1) is a Fourier series.

2. In this section we collect some known results concerning the problems under consideration. The following result is due to Hardy and Littlewood [1]:

THEOREM 1. *If*

$$\sum_1^\infty a_n \cos nt \sim \phi(t) \tag{2.1}$$

is a Fourier series, and if

$$na_n > -C, \qquad n = 1, 2, 3, \cdots; \tag{2.2}$$

C a positive constant, then the necessary and sufficient condition that $\sum a_n$ converges to s, is that

$$1/t \int_0^t \phi(u)\,du \to s, \text{ as } t \downarrow 0. \tag{2.3}$$

In other words: if (2.1) and (2.2) hold, then L-summability implies convergence, and conversely. The assumption (2.1) is equivalent to saying that $F(t)$ in (1.1) is absolutely continuous.

I have generalized this result [3], by replacing (2.2) by the more general assumption:

$$\lim_{\lambda\downarrow 1} \limsup_{n\to\infty} \sum_n^{\lambda n} (|a_\nu| - a_\nu) = 0;^{2} \tag{2.4}$$

i.e. the series $\sum(|a_n| - a_n)$ is slowly oscillating. We have thus

THEOREM 2. *If* (2.1) *and* (2.4) *hold, then convergence of $\sum a_n$ implies L-summability, and conversely.*

Note that if (2.2) holds, then

$$|a_n| - a_n < 2C/n, \qquad n \geqq 1,$$

hence

$$\sum_n^{\lambda n}(|a_\nu| - a_\nu) < 2C \sum_n^{\lambda n} 1/\nu < 2C(1/n + \log\lambda),$$

and (2.4) follows.

Another result of Hardy and Littlewood is:

---

[2] Slightly more was proved there.

THEOREM 3. [2, theorem 14]. *If*

(2.5) $$a_n = O(n^{-\delta}) \text{ for some } \delta > 0, \text{ as } n \to \infty,$$

and if

(2.6) $$s_n - s \equiv a_1 + a_2 + \cdots + a_n - s = o(1/\log n),$$

as $n \to \infty$, then

(2.7) $$L \sum a_n = s.$$

Note that here $\sum a_n \cos nt$ need not be a Fourier series.

3. In this paper we generalize each of the Theorems 1-3. The main results are:

THEOREM 4. *If*

(3.1) $$\lim_{\lambda \downarrow 1} \liminf_{n\to\infty} \min_{n<k\leqq\lambda n} \sum_n^k a_\nu \geqq 0,$$

and if $L \sum a_n = s$, then $\sum a_n$ converges to s.

Note that

$$2a_\nu \geqq a_\nu - |a_\nu|,$$

hence

$$2 \liminf_{n\to\infty} \min_{n<k\leqq\lambda n} \sum_n^k a_\nu \geqq \liminf_{n\to\infty} \sum_n^{\lambda n} (a_\nu - |a_\nu|).$$

It now follows that (2.4) implies (3.1), i.e. (3.1) is more general than (2.4), and *a fortiori* more general than (2.2).

THEOREM 5. If

(3.2) $$\sum_n^{2n} (|a_\nu| - a_\nu) = O(1) \text{ as } n \to \infty,$$

and if

(3.3) $$\sum_1^n a_\nu = O(1) \text{ as } n \to \infty,$$

then

(3.4) $$t^{-1}F(t) = O(1) \text{ as } t \downarrow 0.$$

THEOREM 6. *Assume again (3.2); suppose further that*

(3.5) $$-\infty < \liminf_{n\to\infty} s_n = l \leqq \limsup_{n\to\infty} s_n = u < +\infty.$$

Writing for any $\lambda > 1$

(3.6) $$\limsup_{n\to\infty} \max_{n\leqq\nu<\mu\leqq\lambda n} |s_\mu - s_\nu| = w(\lambda),$$

evidently $w(\lambda) < \infty$. Suppose

(3.7) $$w(\lambda) = o(\log \lambda) \quad \text{as} \quad \lambda \downarrow 1;$$

then

$$l \leqq \liminf_{t \downarrow 0} t^{-1} F(t) \leqq \limsup_{t \downarrow 0} t^{-1} F(t) \leqq u.$$

We shall see that the proof of Theorem 5 is much simpler than that of Theorem 6; the latter however is essentially sharper. In particular, if $u = l = s$, i. e. if $\sum a_n$ converges, then obviously $w(\lambda) \equiv 0$, and we get:

THEOREM 6′. *If* (3.2) *holds, then the convergence of* $\sum a_n$ *implies L-summability of that series.*

THEOREM 7. *Suppose that for some positive* $\delta < 1$

$$\sum_{n}^{2n} (|a_\nu| - a_\nu) = O(n^{1-\delta}) \text{ as } n \to \infty, \tag{3.8}$$

and

$$-\infty < \liminf_{n \to \infty} (s_n - s) \log n = \rho \tag{3.9}$$

$$\leqq \limsup_{n \to \infty} (s_n - s) \log n = \eta < +\infty;$$

then

$$\tfrac{1}{2}(\rho - \eta) \log \delta^{-1} \leqq \liminf_{t \downarrow 0} [t^{-1} F(t) - s] \tag{3.10}$$

$$\leqq \limsup_{t \downarrow 0} [t^{-1} F(t) - s] \leqq \tfrac{1}{2}(\eta - \rho) \log \delta^{-1}.$$

(3.8) is satisfied if in particular

$$|a_n| - a_n = O(n^{-\delta}) \text{ as } n \to \infty,$$

which generalizes (2.5) to a "one-sided" condition. Moreover, if $\rho = \eta$ then we get L-summability, thus generalizing Theorem 3.

Another interesting case occurs when (3.8) holds for *every* $\delta < 1$; for then, letting in (3.10) $\delta \uparrow 1$, we get

THEOREM 7′. *If*

$$\sum_{n}^{2n} (|a_\nu| - a_\nu) = O(n^{1-\delta})$$

for every positive $\delta < 1$, *and if*

$$s_n - s = O(1/\log n)$$

as $n \to \infty$, *then* $L \sum a_n = s$.

This is in a sense the "reciprocal" or "transform" of the following theorem due to Hardy and Littlewood [2, theorem 7]:

If $\phi(t) \sim \sum a_n \cos nt$ is a Fourier series, and if $a_n = O(n^{-\delta})$ for every positive $\delta < 1$, if moreover $\phi(t) = O(1/\log t^{-1})$ as $t \downarrow 0$; then $\sum a_n$ converges. The following is a generalization of Theorem 3 in another direction:

THEOREM 8. *Suppose that for some positive* $\delta < 1$

$$\sum_n^{2n} | a_\nu | = O(n^{1-\delta}) \text{ as } n \to \infty, \tag{3.11}$$

and that there exist two numbers s *and* ω *such that*

$$\sum_{\nu=1}^{n} | s_\nu - s - [\omega/\log(\nu+1)] | = o(n/\log n) \text{ as } n \to \infty; \tag{3.12}$$

then $L \sum a_n = s$.

Again, if (3. 11) holds for *every* $\delta < 1$, then in (3. 12) o can be replaced by O.

4. For the proof of our theorems we introduce some lemmas.

LEMMA 1. *Suppose that* $0 < \delta \leqq 1$,

$$\sum_n^{2n} (| a_\nu | - a_\nu) = O(n^{1-\delta}) \text{ as } n \to \infty, \tag{4.1}$$

and

$$s_n = O(n^{1-\delta}) \text{ as } n \to \infty, \tag{4.2}$$

then

$$\sum_1^\infty \nu^{-1} | a_\nu |$$

converges, and

$$\sum_n^\infty \nu^{-1} | a_\nu | = O(n^{-\delta}), \text{ as } n \to \infty. \tag{4.3}$$

Proof. From (4. 1) and (4. 2)

$$\sum_n^{2n} | a_\nu | = \sum_n^{2n} a_\nu + O(n^{1-\delta}) = O(n^{1-\delta}), \text{ as } n \to \infty, \tag{4.4}$$

hence

$$\sum_n^{2n} \nu^{-1} | a_\nu | \leqq n^{-1} \sum_n^{2n} | a_\nu | = O(n^{-\delta}).$$

Thus for any $k \geqq 0$,

$$\sum_{2^k}^{2^{k+1}-1} \nu^{-1} | a_\nu | = O(2^{-k\delta}), \text{ as } k \to \infty,$$

and summation over k yields

$$\sum_1^{2^l} \nu^{-1} | a_\nu | = O \left| \sum_{k=0}^{l} 2^{-k\delta} \right| = O(1) \text{ as } l \to \infty.$$

Hence $\sum \nu^{-1} \mid a_\nu \mid$ converges, and

$$\sum_{\infty}^{n} \nu^{-1} \mid a_\nu \mid = \sum_{k=0}^{\infty} \sum_{n.2^k}^{n.2^{k+1}-1} \nu^{-1} \mid a_\nu \mid = O(n^{-\delta} \sum_{k=0}^{\infty} 2^{-k\delta}) = O(n^{-\delta}), \text{ as } n \to \infty.$$

This proves the lemma.

Lemma 2. *If* $\sum \nu^{-1} \mid a_\nu \mid < \infty$, *and if*

$$\sum_{n}^{\infty} \nu^{-1} \mid a_\nu \mid = O(n^{-1}), \textit{ as } n \to \infty, \tag{4.5}$$

then

$$\sum_{1}^{n} \nu \mid a_\nu \mid = O(n) \textit{ as } n \to \infty, \tag{4.6}$$

and conversely.

For the proof let

$$\sum_{n}^{\infty} \nu^{-1} \mid a_\nu \mid = r_n, \quad n \geqq 1;$$

then

$$\mid a_n \mid = n(r_n - r_{n+1}), \quad n \geqq 1,$$

and

$$\sum_{1}^{n} \nu \mid a_\nu \mid = r_1 + \sum_{2}^{n} (2\nu - 1) r_\nu - n^2 r_{n+1}, \quad n \geqq 2.$$

By assumption $r_n = O(n^{-1})$, and the last formula yields the first part of the lemma. Conversely, if (4.6) holds, let

$$\sum_{1}^{n} \nu \mid a_\nu \mid = v_n,$$

then $v_n = O(n)$ as $n \to \infty$. Also $n \mid a_n \mid = v_n - v_{n-1}$ for $n \geqq 2$, hence

$$\sum_{2}^{n} \nu^{-1} \mid a_\nu \mid = \sum_{2}^{n} \nu^{-2} v_\nu - \sum_{1}^{n-1} (\nu + 1)^{-2} v_\nu < n^{-2} v_n + \sum_{2}^{n} [\nu^{-2} - (\nu + 1)^{-2}] v_\nu$$
$$= O(n^{-1}) + O(\sum_{1}^{n} \nu^{-2}) = O(1).$$

Thus $\sum \nu^{-1} \mid a_\nu \mid$ converges; furthermore

$$\sum_{n}^{\infty} \nu^{-1} \mid a_\nu \mid = \sum_{n}^{\infty} \nu^{-2} v_\nu - \sum_{n-1}^{\infty} (\nu + 1)^{-2} v_\nu < \sum_{n}^{\infty} [\nu^{-2} - (\nu + 1)^{-2}] v_\nu$$
$$= O(\sum_{n}^{\infty} \nu^{-2}) = O(n^{-1}).$$

This proves the lemma.

Lemma 3. *Let* $\lambda > 1$, *and*

$$\bar{s}_n = \bar{s}_n(\lambda) = \max_{n \leqq \nu < \mu \leqq \lambda n} \mid s_\mu - s_\nu \mid < \gamma \quad \text{for } n \geqq N; \tag{4.7}$$

let $N \leqq u < v$, *and suppose that the sequence*

$$\frac{\sin \nu t}{\nu t},\ \nu = u + 1, u + 2, \cdots, v$$

is monotonic; then

$$\left| \sum_{u+1}^{v} a_\nu \frac{\sin \nu t}{\nu t} \right| < \frac{3\gamma}{\log \lambda} \log \frac{v}{u}.$$

For the proof write

$$\sum_{\nu=u+1}^{v} = \sum_{u<\nu}^{\leqq \lambda u} + \sum_{\lambda u<\nu}^{\leqq \lambda^2 u} + \cdots + \sum_{\lambda^{\rho-1}u<\nu}^{\leqq \lambda^\rho u} + \sum_{\lambda^\rho u<\nu}^{\leqq v}$$

where ρ is the smallest exponent such that

$$v \leqq \lambda^{\rho+1}u, \text{ i.e. } u\lambda^\rho < v \leqq u\lambda^{\rho+1}, \text{ or } \rho < \frac{\log v/u}{\log \lambda} \leqq \rho + 1. \tag{4.8}$$

To each sum apply Abel's summation by part; thus if

$$\frac{\sin \nu t}{\nu t} - \frac{\sin (\nu+1)t}{(\nu+1)t} = \Delta \frac{\sin \nu t}{\nu t},$$

$$\sum_{b+1}^{c} a_\nu \frac{\sin \nu t}{\nu t} = \sum_{b+1}^{c-1} (s_\nu - s_b) \Delta \frac{\sin \nu t}{\nu t} + (s_c - s_b) \frac{\sin ct}{ct}.$$

Here $N \leqq b < c \leqq \lambda b$, and (4.7) yields

$$\left| \sum_{b+1}^{c} a_\nu \frac{\sin \nu t}{\nu t} \right| < \gamma \left| \frac{\sin (b+1)t}{(b+1)t} - \frac{\sin ct}{ct} \right| + \gamma \left| \frac{\sin ct}{ct} \right| < 3\gamma,$$

hence, in view of (4.8)

$$\left| \sum_{u+1}^{v} a_\nu \frac{\sin \nu t}{\nu t} \right| < 3\rho\gamma < \frac{3\gamma}{\log \lambda} \log \frac{v}{u}.$$

LEMMA 4. *For* $0 < (\nu+1)t < \pi/2$,

$$0 < \Delta \frac{\sin \nu t}{\nu t} = \frac{\sin \nu t}{\nu t} - \frac{\sin (\nu+1)t}{(\nu+1)t} < \frac{\nu t^2}{2}; \tag{4.9}$$

and if $\nu t > 1$, *then*

$$\left| \Delta \frac{\sin \nu t}{\nu t} \right| < \frac{1}{\nu} + \frac{1}{\nu(\nu+1)t} < \frac{2}{\nu}. \tag{4.10}$$

Proof. We have obviously

$$\Delta \frac{\sin \nu t}{\nu t} = \int_{\nu t}^{(\nu+1)t} \frac{\sin u - u \cos u}{u^2} du;$$

but from the mean-value theorem

$$0 < \frac{\sin u - u \cos u}{u^2} < (1/3)u \text{ for } 0 < u < \pi/2,$$

hence

$$0 < \Delta \frac{\sin \nu t}{\nu t} < (1/3) \int_{\nu t}^{(\nu+1)t} u \, du = (1/6)(2\nu + 1)t^2 < \nu t^2/2;$$

for $0 < (\nu + 1)t < \pi/2$. Furthermore for $\nu t > 1$,

$$\Delta \frac{\sin \nu t}{\nu t} = \frac{\sin \nu t - \sin (\nu + 1)t}{\nu t} + \frac{\sin (\nu + 1)t}{\nu(\nu + 1)t}$$

$$= \frac{\sin (\nu + 1)t}{\nu(\nu + 1)t} - 2 \frac{\cos (2\nu + 1)t/2 \sin t/2}{\nu t},$$

hence

$$\left| \Delta \frac{\sin \nu t}{\nu t} \right| < \frac{1}{\nu(\nu + 1)t} + \frac{1}{\nu}.$$

5. We proceed to prove the announced results.

Proof of Theorem 4. It is known [Zygmund, 5], that L-summability implies Cesàro-summability of some order; furthermore by a known Tauberian-type theorem Cesàro-summability and (3.1) imply convergence.[3] This proves Theorem 4.

Proof of Theorem 5. From (3.2) and (3.3)

$$\sum_{n}^{2n} | a_\nu | = O(1) + \sum_{n}^{2n} a_\nu = O(1) + s_{2n} - s_{n-1} = O(1), \quad n \to \infty.$$

It now follows from Lemma 1 that $\sum_{n}^{\infty} \nu^{-1} | a_\nu | = O(n^{-1})$, and from Lemma 2, that $\sum_{1}^{n} \nu | a_\nu | = O(n)$.

Now put

$$t^{-1}F(t) = \sum_{1}^{n} a_\nu + \sum_{1}^{n} a_\nu \left(\frac{\sin \nu t}{\nu t} - 1 \right) + \sum_{n+1}^{\infty} a_\nu \frac{\sin \nu t}{\nu t}$$

$$= S_1 + S_2 + S_3;$$

from (3.3) $S_1 = O(1)$ as $n \to \infty$.

From

$$(5.1) \qquad 0 < 1 - \frac{\sin \nu t}{\nu t} < (1/6) \nu^2 t^2,$$

$$| S_2 | \leqq (1/6) t^2 \sum_{1}^{n} \nu^2 | a_\nu | \leqq (1/6) n t^2 \sum_{1}^{n} \nu | a_\nu | = t^2 O(n^2).$$

Furthermore

$$| S_3 | \leqq t^{-1} \sum_{n+1}^{\infty} | a_\nu | / \nu = t^{-1} O(n^{-1});$$

on choosing $n = [t^{-1}]$ for $0 < t < 1$, our theorem follows.

---

[3] A more general theorem has been proved by R. Schmidt in 1925.

Remark. We can replace the assumption (3.3) in Theorem 5 by the more general one:

$$\sum_1^\infty a_\nu r^\nu = O(1) \text{ as } r \uparrow 1. \tag{5.2}$$

We need only show that (3.2) and (5.2) imply (3.3); but this (and somewhat more) has been proved in Lemma 2 of my paper [3]. Note also that conversely (3.3) always implies (5.2); in fact it implies:

$$\sum_1^n a_\nu r^\nu = O(1) \text{ uniformly for } 0 < r < 1 \text{ and for all } n \geqq 1.$$

This follows from

$$\sum_1^n a_\nu r^\nu = (1-r) \sum_1^{n-1} s_\nu r^\nu + s_n r^n,$$

and

$$\Big| \sum_1^n a_\nu r^\nu \Big| \leqq (1-r) \sum_1^{n-1} | s_\nu | r^\nu + | s_n | r^n \leqq 2 \sup | s_\nu | .$$

Proof of Theorem 6. Subtraction of the term $\frac{1}{2}(u+l) \sin t$ from (1.1) shows that we can assume without loss of generality: $l = -u \leqq 0$; we shall prove then that

$$\liminf_{t \downarrow 0} t^{-1} F(t) \geqq -u.$$

We write

$$\begin{aligned}\sum_1^\infty a_\nu \frac{\sin \nu t}{\nu t} &= \sum_1^n a_\nu + \sum_1^n a_\nu \left(\frac{\sin \nu t}{\nu t} - 1 \right) + \sum_{n+1}^{\kappa_n - 1} a_\nu \frac{\sin \nu t}{\nu t} + \sum_{\kappa_n}^\infty a_\nu \frac{\sin \nu t}{\nu t} \\ &\equiv Z_0 + Z_1 + Z_2 + Z_3,\end{aligned}$$

where n and $\kappa_n > n + 1$ will be fixed presently.

To a given positive $\epsilon < 1$ choose $N_1 = N_1(\epsilon) > 2$ so that

$$-u - \epsilon < Z_0 < u + \epsilon \text{ for } n \geqq N_1; \tag{5.3}$$

by (3.7) we can choose $\lambda = \lambda(\epsilon)$ and $N_2 > N_1$ so that

$$\zeta_n \equiv \max_{n \leqq \nu < \mu \leqq \lambda n} | s_\mu - s_\nu | < \epsilon^2 \log \lambda \text{ for } n \geqq N_2.$$

From (5.1) and lemma 2

$$| Z_1 | \leqq (1/6) t^2 \sum_1^n \nu^2 | a_\nu | \leqq (n/6) t^2 \sum_1^n \nu | a_\nu | = t^2 O(n^2);$$

similarly,

$$| Z_3 | \leqq t^{-1} \sum_{\kappa_n}^\infty \nu^{-1} | a_\nu | = t^{-1} O(\kappa_n^{-1}), \text{ as } n \to \infty.$$

We now choose a positive $t < \epsilon/N_2$, and put $n = [\epsilon t^{-1}]$, so that

$$N_2 - 1 < \epsilon t^{-1} - 1 < n \leqq \epsilon t^{-1}.$$

Furthermore put $\kappa_n = [2\epsilon^{-1}t^{-1}]$, so that

$$n \leqq \epsilon t^{-1} < 2\epsilon^{-1}t^{-1} - 2 < \kappa_n - 1 \leqq 2\epsilon^{-1}t^{-1};$$

it then follows that

$$Z_1 = \epsilon O(1) \text{ and } Z_3 = \epsilon O(1) \text{ as } t \downarrow 0. \tag{5.4}$$

It remains to estimate Z_2; subdivide the range $n < \nu < \kappa_n$ into consecutive parts in each of which $(\sin \nu t)/\nu t$ is monotonic. The function $(\sin x)/x$ is monotonic between two consecutive roots of the equation $x \cos x - \sin x = 0$; denoting the positive roots by x_1 x_2, $x_3, \cdots$, we have $x_\nu = \nu\pi + \alpha_\nu$, where $0 < \alpha_\nu < \pi/2$. The sections for the index ν are

$$[\epsilon t^{-1}] < \nu \leqq [x_1 t^{-1}], \quad [x_1 t^{-1}] < \nu \leqq [x_2 t^{-1}], \cdots.$$

Denote the number of these sections by p, then p is the smallest integer such that $\kappa_n - 1 \leqq [x_p t^{-1}]$, thus $[x_{p-1}t^{-1}] < \kappa_n - 1$, and $x_{p-1}t^{-1} < \kappa_n \leqq 2\,\epsilon^{-1}t^{-1}$. This yields

$$(p-1)\pi < 2\epsilon^{-1}, \qquad p - 1 < 2/\pi \cdot \epsilon^{-1}, \qquad p < 1 + 2/\pi \cdot \epsilon^{-1} < 2\epsilon^{-1}.$$

We next estimate the corresponding partial sums:

$$S_{n,1} \equiv \sum_{[\epsilon t^{-1}] < \nu}^{\leqq x_1 t^{-1}}, \qquad S_{n,2} \equiv \sum_{[x_1 t^{-1}] < \nu}^{\leqq x_2 t^{-1}}, \cdots.$$

To this end we use Lemma 3, which yields with $\gamma = \epsilon^2 \log \lambda$,

$$|S_{n,1}| < 3\epsilon^2 \log \frac{[x_1 t^{-1}]}{[\epsilon t^{-1}]} < 3\epsilon^2 \log \frac{x_1 t^{-1}}{\epsilon t^{-1} - 1} < 3\epsilon^2 \log \frac{2x_1}{\epsilon} < 3\epsilon^2 \log 3\pi/\epsilon$$

Similarly

$$|S_{n,2}| < 3\epsilon^2 \log \frac{[x_2 t^{-1}]}{[x_1 t^{-1}]} < 3\epsilon^2 \log \frac{x_2 t^{-1}}{x_1 t^{-1} - 1} < 3\epsilon^2 \log \frac{2x_2}{x_1} < 3\epsilon^2 \log 5;$$

and the same estimate holds for $S_{n,3}, \cdots S_{n,p}$. Thus, as $p < 2\epsilon^{-1}$,

$$|Z_2| < 3\epsilon^2 \{\log 3\pi/\epsilon + 2\epsilon^{-1} \log 5\} < 3\epsilon \{3\pi + 2 \log 5\}. \tag{5.5}$$

Collecting the estimates (5.3), (5.4), and (5.5), we get

$$\liminf_{t \downarrow 0} t^{-1}F(t) \geqq -u - \epsilon O(1);$$

but ϵ is arbitrarily small, hence

$$\liminf_{t \downarrow 0} t^{-1}F(t) \geqq -u,$$

and similarly

$$\limsup_{t \downarrow 0} t^{-1}F(t) \leqq u.$$

This proves Theorem 6. This in turn yielded Theorem 6′, a generalization of which is:

THEOREM 6″. *If* (3.2) *and* (3.1) *hold, then the Abel-summability of* $\sum a_n$ *implies its L-summability.*

By a theorem of R. Schmidt, Abel-summability and (3.1) imply convergence, and Theorem 6″ now follows from 6′.

Note also that (3.1) is more general than (3.7).

6. *Proof of Theorem 7. We introduce the function*

$$h(t) = \sum_1^\infty \frac{\alpha_n}{n} \sin nt \equiv (1 - (1/\log 3))\frac{\sin 2t}{2} + \sum_3^\infty \left(\frac{1}{\log n} - \frac{1}{\log (n+1)}\right)\frac{\sin nt}{n},$$

so that

$$\alpha_1 = 0, \quad \alpha_2 = 1 - (1/\log 3), \quad \alpha_n = \frac{1}{\log n} - \frac{1}{\log (n+1)} \text{ for } n \geqq 3.$$

Let for $n \geqq 1$

$$A_n = \sum_1^n \alpha_\nu = 1 - \frac{1}{\log (n+1)};$$

hence

(6.1) $$A_n \to 1, \text{ and } (A_n - 1) \log n \to -1, \text{ as } n \to \infty.$$

Furthermore

(6.2) $$\alpha_n = \frac{\log (1 + 1/n)}{\log n \cdot \log (n+1)} \sim \frac{1}{n \log^2 n}, \text{ as } n \to \infty;$$

and from $\sum_1^\infty \alpha_n < \infty$ it is easily seen that

(6.3) $$t^{-1}h(t) \to 1 \text{ as } t \downarrow 0.$$

We write

$$\tfrac{1}{2}(\eta + \rho) = \beta, \quad s + \beta = s', \quad a_n + \beta\alpha_n = a'_n, \quad \sum_1^n a'_\nu = s'_n;$$

then from (3.8) and (6.2)

(6.4) $$\sum_n^{2n} (|a'_\nu| - a'_\nu) = O(n^{1-\delta}) + O\left(\sum_n^{2n} \frac{1}{\nu \log^2 \nu}\right)$$
$$= O(n^{1-\delta}) + O\left(\frac{1}{\log^2 n}\right) = O(n^{1-\delta}), \text{ as } n \to \infty.$$

Furthermore from (3.9) and (6.1)

(6.5) $$\limsup_{n\to\infty} (s'_n - s') \log n = \eta - \beta = (\eta - \rho)/2 \equiv \sigma,$$

and

(6.6) $$\liminf_{n\to\infty} (s'_n - s') \log n = \rho - \beta = -\sigma.$$

We now consider

$$F(t) + \beta h(t) = \sum_1^\infty (a'_\nu/\nu) \sin \nu t,$$

and write

$$\sum_1^\infty a'_\nu (\sin \nu t)/\nu t - s' = (\sum_1^n a'_\nu (\sin \nu t)/\nu t - s') + \sum_{n+1}^{\kappa_n} + \sum_{1+\kappa_n}^{\infty} \equiv U_1 + U_2 + U_3.$$

We choose an r such that

$$r\delta > 1, \quad \text{i.e.} \quad r > \delta^{-1};$$

to a given $t < \frac{1}{4}$ choose

(6.7) $$n = [t^{-1}], \text{ hence } 1 < t^{-1} - 1 < n \leqq t^{-1},$$

and let

(6.8) $$\kappa_n = [t^{-r}], \text{ hence } n + 1 \leqq 1 + t^{-1} < t^{-r} - 1 < \kappa_n \leqq t^{-r},$$

for $t < t(\delta) < \frac{1}{4}$. Now

(6.9) $$U_1 = \sum_1^{n-1} (s'_\nu - s') \Delta (\sin \nu t)/\nu t + (s'_n - s')(\sin nt)/nt + s'((\sin t)/t - 1),$$

and from (3.9) and Lemma 4,

$$U_1 = O(1) \sum_2^n \frac{\nu t^2}{\log \nu} + O\left(\frac{1}{\log n}\right) + o(1);$$

thus from (6.7)

(6.10) $$U_1 = O(1) n^{-1} \sum_2^n \frac{1}{\log \nu} + o(1) = o(1) \text{ as } t \downarrow 0.$$

Next from (6.4), (6.5) and (6.6) by Lemma 1,

(6.11) $$| U_3 | < t^{-1} \sum_{\kappa_n}^\infty \nu^{-1} | a'_\nu | = t^{-1} O(\kappa_n^{-\delta}) = O(t^{r\delta-1}) = o(1) \text{ as } t \downarrow 0.$$

Finally

(6.12) $$U_2 = \sum_{n+1}^{\kappa_n-1} (s'_\nu - s') \Delta (\sin \nu t)/\nu t + (s'_{\kappa_n} - s')(\sin \kappa_n t)/\kappa_n t,$$

and, using (4.10), (6.5) and (6..6) we get for $t < t_0(\epsilon)$

$$| U_2 | < (\sigma + \epsilon) t^{-1} \sum_{n+1}^{\kappa_n-1} \frac{1}{\nu(\nu+1) \log \nu} + (\sigma + \epsilon) \left(\sum_{n+1}^{\kappa_n} \frac{1}{\nu \log \nu} + \frac{1}{\log \kappa_n}\right);$$

thus from (6. 7) and (6. 8)

$$(6.13)\quad |U_2| < \frac{(\sigma+\epsilon)(n+1)}{\log(n+1)}\left(\frac{1}{n+1}-\frac{1}{\kappa_n}\right)+\frac{2(\sigma+\epsilon)}{\log(n+1)} + (\sigma+\epsilon)\int_{n+1}^{\kappa_n}\frac{dx}{x\log x}$$

$$< \frac{3(\sigma+\epsilon)}{\log t^{-1}} + (\sigma+\epsilon)(\log\log t^{-r} - \log\log t^{-1})$$

$$< \frac{3(\sigma+\epsilon)}{\log t^{-1}} + (\sigma+\epsilon)\log r.$$

Hence, ϵ being arbitrarily small,

$$\limsup_{t\downarrow 0} |U_2| \leqq \sigma \log r;$$

but r is arbitrarliy close to δ^{-1} so that

$$\limsup_{t\downarrow 0} |U_2| \leqq \sigma\log\delta^{-1} = \tfrac{1}{2}(\eta-\rho)\log\delta^{-1},$$

and, in view of (6. 10) and (6. 11)

$$\limsup_{t\downarrow 0} \left|\sum_1^\infty a'_\nu(\sin\nu t)/\nu t - s'\right| \leqq \tfrac{1}{2}(\eta-\rho)\log\delta^{-1},$$

or

$$\limsup_{t\downarrow 0} |t^{-1}F(t) - s + \beta t^{-1}h(t) - \beta| \leqq \tfrac{1}{2}(\eta-\rho)\log\delta^{-1}.$$

Now (6. 3) yields

$$-\tfrac{1}{2}(\eta-\rho)\log\delta^{-1} \leqq \liminf_{t\downarrow 0}(t^{-1}F(t)-s) \leqq \limsup_{t\downarrow 0}(t^{-1}F(t)-s) \leqq \tfrac{1}{2}(\eta-\rho)\log\delta^{-1}.$$

This proves Theorem 7.

7. We proceed to prove Theorem 8. It follows from (3. 11) and (3. 12) that $s_n = O(n^{1-\delta})$, and from Lemma 1 that $\sum_1^\infty \nu^{-1}|a_\nu| < \infty$, and that (4. 3) holds. We write as in Section **6**.

$$\sum_1^\infty a'_\nu(\sin\nu t)/\nu t - s' = \left(\sum_1^n a'_\nu(\sin\nu t)/\nu t - s'\right) + \sum_{n+1}^{\kappa_n} + \sum_{1+\kappa_n}^\infty \equiv U_1 + U_2 + U_3,$$

where now, using (6. 2)

$$a'_n = a_n + \omega\alpha_n,\qquad s'_n = s_n + \omega - \omega/\log(n+1),\qquad s' = s+\omega.$$

In the first place from (6.9) and (5.1)

$$|U_1| \leqq \sum_1^{n-1} |s'_\nu - s'| \, |\Delta(\sin \nu t)/\nu t| + |s'_n - s'| + |s'| \, t^2;$$

now (4.9) yields

$$|U_1| \leqq |s'_n - s'| + |s'| \, t^2 + \sum_1^{n-1} |s'_\nu - s'| \, \nu t^2$$
$$\leqq |s'| \, t^2 + |s'_n - s'| + nt^2 \sum_1^{n} |s'_\nu - s'| .$$

But from (3.12)

$$\sigma_n \equiv \sum_1^{n} |s'_\nu - s'| = o(n/\log n) \text{ as } n \to \infty, \tag{7.1}$$

hence

$$|U_1| < |s'_n - s'| + |s'| \, t^2 + t^2 o(n^2/\log n) \text{ as } n \to \infty. \tag{7.2}$$

To a given $\epsilon > 0$ choose $N(\epsilon)$ such that

$$\sum_1^{n} |s'_\nu - s'| < \epsilon(n/\log n) \text{ for } n > N(\epsilon).$$

Denote by v_n the number of those indices $\nu \leqq n$ for which

$$|s'_\nu - s'| \geqq \epsilon,$$

then

$$\epsilon(n/\log n) > \sum_1^{n} |s'_\nu - s'| \geqq \epsilon v_n,$$

hence

$$v_n < (n/\log n).$$

Thus: in the interval $1, 2, \cdots, n$ there are at least $n(1 - 1/\log n)$ indices for which $|s'_\nu - s'| < \epsilon$. Choosing now $t < t_0(\epsilon)$, we can always find an n in the interval $t^{-1} < n < 2t^{-1}$ for which $|s'_n - s'| < \epsilon$, and from (7.2)

$$|U_1| < \epsilon + o(1), \text{ as } t \downarrow 0. \tag{7.3}$$

Next from (6.12) and (4.10)

$$|U_2| \leqq 2 \sum_{n+1}^{\kappa_n} |s'_\nu - s'| \, \nu^{-1} + |s'_{\kappa_n} - s'| \, 1/\kappa_n t.$$

We now choose an r such that $r\delta > 1$, and a $\kappa_n > n + 1$ in the interval $t^{-r} < \kappa_n < 2t^{-r}$, such that $|s'_{\kappa_n} - s'| < \epsilon$. Then

$$1/\kappa_n t \, |s'_{\kappa_n} - s'| < \epsilon \, O(t^{r-1}) = o(1).$$

Also, from (7.1)

$$\sum_{n+1}^{\kappa_n} \nu^{-1} \mid s'_\nu - s' \mid = \sum_{n+1}^{\kappa_n - 1} \frac{\sigma_\nu}{\nu(\nu+1)} - \frac{\sigma_n}{n+1} + \frac{\sigma_{\kappa_n}}{\kappa_n}$$

$$= o\Big(\sum_{n+1}^{\kappa_n} 1/\nu \log \nu\Big) + o(1) = o(1),$$

(Cf. the estimate in (6.13)). Thus

$$U_2 = o(1) \text{ as } t \downarrow 0. \tag{7.4}$$

Finally, as in (6.11)

$$\mid U_3 \mid \leqq t^{-1} \sum_{\kappa_n}^{\infty} \nu^{-1} \mid a'_\nu \mid = o(1) \text{ as } t \downarrow 0. \tag{7.5}$$

Summarizing, (7.3), (7.4), and (7.5) yield, letting $\epsilon \downarrow 0$, Theorem 8. If we replace (3.11) by the sharper assumption:

$$\sum_{n}^{2n} \mid a_\nu \mid = O(n^{1-\delta}) \text{ for } \textit{every} \text{ positive } \delta < 1, \text{ as } n \to \infty, \tag{7.6}$$

and (3.12) by the weaker assumption

$$\sum_{1}^{n} \mid s_\nu - s \mid = O(n/\log n) \text{ as } n \to \infty, \tag{7.7}$$

then write

$$\sum_{1}^{\infty} a_\nu(\sin \nu t)/\nu t - s = \Big(\sum_{1}^{n} a_\nu(\sin \nu t)/\nu t - s\Big)$$

$$+ \sum_{n+1}^{\kappa_n} + \sum_{1+\kappa_n}^{\infty} \equiv V_1 + V_2 + V_3.$$

If i_n is the number of those indices $\nu \leqq n$ for which $\mid s_\nu - s \mid \geqq \epsilon$, then by (7.7) for a constant γ

$$\gamma(n/\log n) > \sum_{1}^{n} \mid s_\nu - s \mid \geqq \epsilon i_n,$$

or,

$$i_n < \epsilon^{-1}\gamma(n/\log n) < n/3 \text{ for } n > N(\epsilon).$$

Thus: for any $n > N$ in the interval $1, 2, \cdots, n$ there are at least $[2n/3]$ indices for which $\mid s_\nu - s \mid < \epsilon$. We then find as in (7.2) with an n in $t^{-1} < n < 2t^{-1}$,

$$\mid V_1 \mid < \epsilon + O(t^2) + O(n^2t^2/\log n) < \epsilon + o(1) \text{ as } t \downarrow 0.$$

Similarly

$$| V_2 | < \epsilon O(t^{r-1}) + O(1/\log \kappa_n) + O(\sum_{n+1}^{\kappa_n} 1/\nu \log \nu)$$

$$< \epsilon O(1) + o(1) + O(\int_n^{\kappa_n} dx/x \log x)$$

$$< \epsilon O(1) + o(1) + O(1) \log r,$$

and

$$V_3 = t^{-1} O(\kappa_n^{-\delta}) = O(t^{r\delta - 1}).$$

We now choose $r = 1 + \epsilon$, and $(1 + \epsilon)\delta > 1$, i. e. $\delta > 1/(1 + \epsilon)$. Then as $t \downarrow 0$,

$$\limsup | \sum_1^\infty a_\nu (\sin \nu t)/\nu t - s | \leqq \epsilon O(1) + O(1) \log (1 + \epsilon).$$

This gives, letting $\epsilon \downarrow 0$,

THEOREM 8′. *If* (7.6) *and* (7.7) *hold, then* $L \sum a_n = s$. *This is an analogue of Theorem* 7′.

(Added March 1942.)

In Theorem 8 we can replace assumption (3.11) by assumption (3.8). We need only to show that (3.8) and (3.12) imply $s_n = O(n^{1-\delta})$. We have from (3.8)

$$\sum_1^n \nu(| a_\nu | - a_\nu) = \sum_{k=0}^{n} \sum_{2^{-k-1}n+1}^{2^{-k}n} \nu(|a_\nu | - a_\nu) \leqq \sum_k (2^{-k}n \sum_{2^{-k-1}n+1}^{2^{-k}n} (| a_\nu | - a_\nu) = O(n^{2-\delta} \sum 2^{-\kappa}$$

$$= O(n^{2-\delta}),$$

and from (3.12)

$$\sum_1^n | s_\nu | = O(n), \text{ as } n \to \infty.$$

Put $n\sigma_n = \sum_1^n s_\nu$, then $\sigma_n = O(1)$, and

$$n\sigma_n - (n + 1)s_n = - (a_1 + 2a_2 + \cdots + na_n) \leqq \tfrac{1}{2} \sum_1^n \nu(| a_\nu | - a_\nu)$$

$$< c_1 n^{2-\delta}, \text{ as } n \to \infty,^4$$

hence

$$(n + 1)s_n > - 2c_1 n^{2-\delta}.$$

Furthermore

$$s_n = \sigma_n - 2(\sigma_{2n} - \sigma_n) - n^{-1} \sum_{\nu=1}^n (s_{n+\nu} - s_n),$$

---

[4] $c_1, c_2, \ldots$ are constants.

and

$$-(s_{n+\nu}-s_n) \leqq \tfrac{1}{2}\sum_{n+1}^{n+\nu}(|a_\kappa|-a_\kappa) < c_2 n^{1-\delta}$$

hence

$$s_n < c_3 n^{1-\delta},$$

which proves $s_n = O(n^{1-\delta})$. The rest of the proof remains unaltered.

Accordingly in Theorem 8′ (7.6) can be replaced by:

$$\sum_{n}^{2n}(|a_\nu|-a_\nu) = O(n^{1-\delta}) \text{ for } \textit{every} \text{ positive } \delta < 1, \text{ as } n \to \infty.$$

University of Cincinnati.

BIBLIOGRAPHY

[1] G. H. Hardy and J. E. Littlewood, "Two theorems concerning Fourier series," *Journal of the London Mathematical Society*, vol. 1 (1926), pp. 19-25.

[2] G. H. Hardy and J. E. Littlewood, "Some new convergence criteria for Fourier series," *Annali della R. Scuola Normale Superiore di Pisa, Serie II*, vol. 3 (1934), pp. 43-62.

[3] O. Szász, "Convergence properties of Fourier series," *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 483-500.

[4] S. Verblunsky, "On the theory of trigonometric series (II)," *Proceedings of the London Mathematical Society, Ser. 2*, vol. 34 (1932), pp. 457-491.

[5] A. Zygmund, "Sur la dérivation des séries de Fourier," *Bulletin International de l'Académie Polonaise des Sciences et des Lettres*, Serie A (1924), pp. 243-249.

[6] A. Zygmund, *Trigonometrical series* (1935).

REPEATED LINEAR FACTORS OF FORMS.*

By Rufus Oldenburger.

---

1. Introduction. We consider a form $\mathcal{F}$ of degree n, given by

$$\mathcal{F} = \sum_1^m a_{ij\ldots k} x_i x_j \cdots x_k,$$

where the a's are elements from a given field $\mathcal{K}$, and are understood to be such that any permutation of the subscripts of an element does not change the value of the element.

With $\mathcal{F}$ is associated the completely symmetric n-way matrix $\mathcal{A} = (a_{ij\ldots k})$, unique if the characteristic of $\mathcal{K}$ does not divide the multinomial coefficients in the expansion of $(x_1 + \cdots + x_m)^n$. It is assumed in this paper that the characteristic of $\mathcal{K}$ has this property. We can arrange the elements of $\mathcal{A}$ in ordinary 2-way matrices $\mathcal{A}_0, \mathcal{A}_1, \cdots, \mathcal{A}_n$, where for each σ the matrix $\mathcal{A}_\sigma$ is obtained by choosing σ indices of $\mathcal{A}$ for the multipartite column index of $\mathcal{A}_\sigma$, and the remaining indices for the multipartite row index of $\mathcal{A}_\sigma$, the values of the indices varying in some prescribed order. Thus the values of σ indices of $\mathcal{A}$ vary with the columns of $\mathcal{A}_\sigma$, and the values of the remaining indices vary with the rows.

The matrices $\mathcal{A}_0, \mathcal{A}_1, \cdots, \mathcal{A}_\sigma$ have been used by Walther Mayer, Thrall and Chandler, Hitchcock, the author, and others. These matrices are of some importance in the theory of forms because their ranks are invariant under non-singular linear transformations on the variables in $\mathcal{F}$. The rank of $\mathcal{A}_1$ has a simple interpretation in that it yields the number of essential variables in $\mathcal{F}$, where it is understood that r is the number of essential variables in $\mathcal{F}$ if $\mathcal{F}$ cannot be brought by a non-singular linear transformation into a form $\mathcal{F}'$ with less than r variables. In the present paper we shall show how these matrices and their ranks are connected in a direct and very simple way with the linear factors of $\mathcal{F}$ and their multiplicities. We shall prove that $\mathcal{F}$ *has a linear factor* $\mathcal{L}$ *repeated* r *times if and only if* $\mathcal{L}^{n-r+1}$ *is a covariant constructible from the matrix* $\mathcal{A}_{n-r+1}$.

2. Matrices and derivatives of $\mathcal{F}$. Except for rational factors all s-th derivatives of $\mathcal{F}$ are of the form

$$\sum_1^m a_{\gamma\alpha\ldots\beta} x_\alpha \cdots x_\beta, \tag{2.1}$$

---

[1] Received June 14, 1941; Revised Sept. 5, 1941; Presented to the American Mathematical Society, September 2, 1941.

where τ is a fixed set of s indices of $\mathfrak{A}$. The matrix of the forms (2.1) is $\mathfrak{A}_{n-s}$. We have the following lemma.

LEMMA 2.1. *The form $\mathfrak{F}$ has the factor $x_1{}^r$ if and only if $\mathfrak{A}_{n-r+1}\,\xi = 0$ has each vector $\xi = (\xi_{\alpha\ldots\beta})$ as a solution where ξ is a vector with $\xi_{\alpha\ldots\beta} = 0$ when at least one subscript is* 1.

Lemma 2.1 follows from the property that if $\mathfrak{F}$ has $x_1{}^r$ as a factor the derivatives (2.1) with $s = r - 1$ vanish identically in $x_2, \cdots, x_m$ when we set $x_1 = 0$, and conversely.

In the following corollary "singular on its columns" means that the rank of $\mathfrak{A}_{n-r+1}$ is not equal to the number of distinct columns in the most general matrix of the type $\mathfrak{A}_{n-r+1}$; that is the column rank is not a maximum.

COROLLARY 2.1. *If $\mathfrak{F}$ has a linear factor repeated at least r times the matrix $\mathfrak{A}_{n-r+1}$ is singular on its columns.*

We consider the equation $\mathfrak{A}_\rho\eta = 0$ where the column vector η has ρ indices $\alpha, \beta, \cdots, \gamma$ varying in the same manner with respect to the rows of η as the column indices of $\mathfrak{A}_\rho$ vary with the columns of $\mathfrak{A}_\rho$. We write η as $(\eta_{\alpha\beta\ldots\gamma})$, and assume that $(\eta_{\alpha\beta\ldots\gamma})$ is completely symmetric when considered as a ρ-way matrix with the indices $\alpha, \beta, \cdots, \gamma$. We introduce the form $\mathfrak{F}_\rho$ given by

$$(2.2) \qquad \sum_1^m \eta_{\alpha\beta\ldots\gamma}\mu_\alpha\mu_\beta \cdots \mu_\gamma.$$

The forms $\{\mathfrak{F}_\rho\}$ obtained by taking the various solutions $\{\eta\}$ of $\mathfrak{A}_\rho\eta = 0$ form a linear space of forms, which space we shall denote by $\mathfrak{S}_\rho$.

The *direct product* of spaces $\mathfrak{D}$ and $\mathfrak{B}$ is obtained by multiplying all elements in $\mathfrak{D}$ by the elements in $\mathfrak{B}$. A space $\mathfrak{D}$ is understood to *contain* a space $\mathfrak{B}$ if $\mathfrak{B}$ is $\mathfrak{D}$ or is a subspace of $\mathfrak{D}$.

If $\mathfrak{F}$ has $x_1{}^r$ as a factor, by Lemma 2.1 the space $\mathfrak{S}_{n-r+1}$ contains the space $\mathfrak{R}$ of all forms in $\mu_2, \cdots, \mu_m$ of degree $n - r + 1$. Since $x_1{}^{r-1}$ is a factor of $\mathfrak{F}$, by the same lemma the space $\mathfrak{T}$ of all forms in $\mu_2, \cdots, \mu_m$ of degree $n - r + 2$ is contained in $\mathfrak{S}_{n-r+2}$. Since $\mathfrak{T}$ contains the direct product of $\mathfrak{R}$ by the space of all linear forms in $\mu_2, \cdots, \mu_m$ the following theorem is suggested.

THEOREM 2.1. *If $\sigma > \rho$ and the space $\mathfrak{S}_\rho$ has dimension at least* 1, *the space $\mathfrak{S}_\sigma$ contains the direct product of the space $\mathfrak{S}_\rho$ and the space of all forms in $\mu_1, \cdots, \mu_m$ of degree $\sigma - \rho$.*

We can write the matrix $\mathfrak{A}_\rho$ as

$$\begin{Vmatrix} \xi_{1\lambda} \cdots \xi_{1\mu} \\ \cdot \quad \cdot \quad \cdot \\ \xi_{m\lambda} \cdots \xi_{m\mu} \end{Vmatrix},$$

where the ξ's are column vectors with $n - \rho - 1$ indices varying with the rows of these ξ's, while $\mathfrak{A}_{\rho+1}$ is given by

$$\| \xi_{1\lambda} \cdots \xi_{1\mu} \cdots \xi_{m\lambda} \cdots \xi_{m\mu} \|.$$

Theorem 2.1 follows from the relation indicated between $\mathfrak{A}_\rho$ and $\mathfrak{A}_{\rho+1}$.

COROLLARY 2.2. *If the matrix $\mathfrak{A}_\rho$ is singular on its columns the matrix $\mathfrak{A}_\sigma$ is singular on its columns for each σ such that $\sigma > \rho$.*

3. **Adjoints and linear factors of forms.** We consider any form written as (2.2) where $(\eta_{\alpha\beta\ldots\gamma})$ is completely symmetric and not zero, and we construct a form $\mathcal{B}_\rho$ of degree ρ given by

$$\mathcal{B}_\rho = \sum_1^m \phi_{\alpha\lambda\ldots\psi} w_\alpha w_\lambda \cdots w_\psi,$$

where the coefficients $(\phi_{\alpha\lambda\ldots\psi})$ in $\mathcal{B}_\rho$ are the elements in a solution $(\phi_{\alpha\lambda\ldots\psi})$ of

$$(\sum_{\alpha=1}^m \eta_{\alpha\beta\ldots\gamma}\phi_{\alpha\lambda\ldots\psi}) = 0, \tag{3.1}$$

and further $(\phi_{\alpha\lambda\ldots\psi})$ is understood to be a completely symmetric ρ-way matrix. If $(\phi_{\alpha\lambda\ldots\psi}) \neq 0$, we say that $\mathcal{B}_\rho$ is an *adjoint* form of the form (2.2).

The forms of type $\mathcal{B}_\rho$ which are adjoints of forms in the space $\mathcal{S}_\rho$ form a space $\mathcal{T}_\rho$ of forms. We term $\mathcal{B}_\rho$ a *ρ-adjoint* of $\mathcal{F}$. Under a non-singular linear transformation bringing $\mathcal{F}$ into a form $\mathcal{F}'$ the forms in $\mathcal{T}_\rho$ transform covariantly into the ρ-adjoint forms of $\mathcal{F}'$. For this reason we write x_i in place of w_i for each i.

In the following theorem "non-vacuous" means "contains a non-zero form."

THEOREM 3.1. *If ρ is such that the space $\mathcal{T}_\rho$ is non-vacuous, for each σ greater than ρ the space $\mathcal{T}_\sigma$ is non-vacuous.*

If $\mathcal{T}_\rho$ is non-vacuous there is a ρ-adjoint $\mathcal{B}_\rho$ in $\mathcal{T}_\rho$ related to a form $\mathcal{F}_\rho$ in $\mathcal{S}_\rho$ by the equation (3.1), where $\mathcal{F}_\rho \not\equiv 0$. For (3.1) to have a non-zero solution $(\phi_{\alpha\lambda\ldots\psi})$ it is necessary that the rank of the matrix $\mathcal{H}_\rho$ be less than m where $\mathcal{H}_\rho$ is an ordinary 2-way display of the symmetric ρ-way matrix $(\eta_{\alpha\beta\ldots\gamma})$ obtained by letting α be the row index of $\mathcal{H}_\rho$ and $\beta, \cdots, \gamma$ the multipartite column index. It follows that $\mathcal{F}_\rho$ is a form in r essential variables, where $r < m$, which can be reduced by a non-singular linear transformation to a form in r variables. We may therefore assume to begin with that $\mathcal{F}_\rho$ is a form in $\mu_1, \cdots \mu_r$. By Theorem 2.1 the form $\mathcal{L}\mathcal{F}_\rho$ is contained in $\mathcal{S}_{\rho+1}$ where $\mathcal{L}$ is a linear form in $\mu_1, \cdots, \mu_r$ not identically zero. If we construct the matrix $\mathcal{H}_{\rho+1}$ of $\mathcal{L}\mathcal{F}_\rho$ which corresponds to $\mathcal{H}_\rho$ for $\mathcal{F}_\rho$ we find that $\mathcal{H}_{\rho+1}$ is at most of rank r. Since $\mathcal{L}\mathcal{F}_\rho \not\equiv 0$ the matrix $\mathcal{H}_{\rho+1}$ is not zero. It follows that we can solve the analogue of (3.1) with $\mathcal{F}_{\rho+1} \equiv \mathcal{L}\mathcal{F}_\rho$ and $\mathcal{S}_{\rho+1}$ replacing $\mathcal{F}_\rho$ and $\mathcal{S}_\rho$ respectively, and obtain an adjoint $\mathcal{S}_{\rho+1}$ of $\mathcal{L}\mathcal{F}_\rho$.

We remark that the linear space of forms $\{\mathcal{S}_\rho\}$ adjoint to a form $\mathcal{F}_\rho$ is a space of dimension $m - r$, and that the linear space of forms $\{\mathcal{S}_{\rho+1}\}$ adjoint to $\mathcal{L}\mathcal{F}_\rho$ is of dimension at least $m - r$.

THEOREM 3.2. *The form $\mathcal{F}$ has $\mathcal{L}$ as a linear factor repeated at least r times if and only if the space $\mathcal{F}_{n-r+1}$ contains the form $\mathcal{L}^{n-r+1}$, and all forms of which $\mathcal{L}^{n-r+1}$ is an adjoint are contained in $\mathcal{S}_{n-r+1}$.*

Theorem 3.2 follows from Lemma 2.1 and the relation (3.1), with $\rho = n - r + 1$, defining $\mathcal{S}_{n-r+1}$.

It is evident that if $\mathcal{L}^\rho$ is contained in $\mathcal{F}_\rho$, the form $\mathcal{L}_\sigma$ is contained in $\mathcal{F}_\sigma$ for $\sigma > \rho$. That $\mathcal{F}_\rho$ exists when $\mathcal{F}_\rho$ exists was already brought out in Theorem 3.1 above.

A form $\mathcal{S}_{n-r+1}$ is the $(n - r + 1)$-st power of a linear form if and only if the analogue of $\mathfrak{A}_1$ for $\mathcal{S}_{n-r+1}$ has rank 1. Thus the *multiplicities* of the linear factors of $\mathcal{F}$ can be expressed in terms of *ranks* of matrices of adjoints of $\mathcal{F}$.

4. Determinants. If $\mathcal{S}_\rho$ is a quadratic the condition (3.1) is one satisfied by the classical adjoint of (2.2), as defined in the literature, when the quadratic form (2.2) is singular. In this section we shall show that we can associate a space of forms with $\mathcal{F}$ defined differently than $\mathcal{F}_{n-r+1}$ such that this space contains the form $\mathcal{L}^{n-r+1}$ if $\mathcal{F}$ has the factor $\mathcal{L}^r$, where $\mathcal{L}$ is linear. Although this definition has the disadvantage that Theorem 3.2 must be modified to obtain a substitute in terms of this space, on the other hand the forms in this space are closely connected with higher dimensional determinants. If the forms in $\mathcal{S}_\rho$ are singular in a certain sense the space $\mathcal{F}_\rho$ is a subspace of a space to be defined here.

We replace condition (3.1) by the condition

$$(\sum_1^m \eta_{\alpha\beta\ldots\gamma}\phi_{\xi\beta\ldots\gamma}) = (\delta_{\alpha\xi}), \tag{4.1}$$

where $(\delta_{\alpha\xi})$ is the Kronecker delta whenever there is a non-zero completely symmetric solution $(\phi_{\xi\beta\ldots\gamma})$ of (4.1) with $(\delta_{\alpha\xi})$ as the Kronecker delta. If there is no completely symmetric solution $(\phi_{\xi\beta\ldots\gamma})$ of (4.1) with $(\delta_{\alpha\xi})$ as the Kronecker delta, it will be understood that $(\phi_{\xi\beta\ldots\gamma})$ is the completely symmetric solution of (4.1) with $(\delta_{\alpha\xi}) = 0$. We shall consider forms $\{\mathcal{S}_\rho\}$ so defined.

We apply to (2.2) the non-singular linear transformation $\mu_\alpha = \Sigma b_{\alpha i} v_i$, and the contragredient transformation to a form $\mathcal{S}_\rho$ defined as in the preceding paragraph. We let the new forms be designated by $\mathcal{R}$ and $\mathcal{S}$ respectively. The matrix for $\mathcal{R}$ and $\mathcal{S}$ corresponding to the left member of (4. 1) for (2. 2) and $\mathcal{B}_\rho$ reduces after some preliminary computation to

$$(\mathcal{M}_{i\alpha}) = (\sum_1^m \eta_{\alpha\beta\ldots\gamma}\phi_{\xi\beta\ldots\gamma}b_{\alpha i}B_{\xi\alpha}),$$

where $(B_{\xi\alpha})$ is the inverse of $(b_{\alpha i})$. By (4. 1) we have

$$(\mathcal{M}_{i\alpha}) = (\delta_{i\alpha}). \tag{4. 2}$$

Thus $\mathcal{B}_\rho$ transforms contravariantly.

We let the form (2. 2) be an arbitrary form of even degree. In what follows we assume that all indices are signant in the higher dimensional determinants involved. We replace each element by its cofactor in $(\eta_{\alpha\beta\ldots\gamma})$. We thus obtain a matrix $(\mathcal{H}_{\xi\beta\ldots\gamma})$. If $\alpha = \xi$ the sum

$$\sum_1^m \eta_{\alpha\beta\ldots\gamma}\mathcal{H}_{\xi\beta\ldots\gamma} \tag{4. 3}$$

is the determinant Δ of $(\eta_{\alpha\beta\ldots\gamma})$. If $\alpha \neq \xi$ the sum (4. 3) vanishes. Thus $(\phi_{\xi\beta\ldots\gamma}) = (\mathcal{H}_{\xi\beta\ldots\gamma})/\Delta$ satisfies (4. 1) when $\Delta \neq 0$. It is readily seen that if $\mathcal{F}$ has the factor $\mathcal{L}^r$, $\mathcal{L}$ linear, the form $\mathcal{L}^{n-r+1}$ belongs to the space of forms

$$\sum_1^m \mathcal{H}_{\xi\beta\ldots\gamma}x_\xi x_\beta \cdots x_\gamma$$

obtained by letting (2. 2) vary over the forms in $\mathcal{S}_{n-r+1}$.

ILLINOIS INSTITUTE OF TECHNOLOGY.

BIBLIOGRAPHY

1. Walther Mayer, "Die Differentialgeometrie der Untermannigfaltigkeiten des $\mathcal{R}_n$ konstanter Krümmung," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 274-310.
2. Thrall and Chandler, "Ternary trilinear forms in the field of complex numbers," *Duke Mathematical Journal*, vol. 4 (1938), pp. 678-690.
3. R. Oldenburger, "Rational equivalence of a form to a sum of p-th powers," *Transactions of the American Mathematical Society*, vol. 44 (1938), pp. 219-249.
4. R. Oldenburger, "Composition and rank of n-way matrices and multilinear forms," *Annals of Mathematics*, vol. 35 (1934), pp. 622-653.
5. F. L. Hitchcock, "Multiple invariants and generalized rank of a p-way matrix or tensor," *Journal of Mathematics and Physics*, vol. 7 (1927), pp. 40-79.
6. R. Oldenburger, "Polynomials in several variables," *Annals of Mathematics*, vol. 41 (1940), pp. 694-710.

SOME PROPERTIES OF BANACH SPACES.*

By D. G. Bourgin.

1. Introduction.[1] This paper is concerned with various topological characterizations of Banach spaces and their subsets [1]. The main concepts are those of the weak topologies, weak compactness and completeness of order $\aleph_a$, the property (Q), weak convergence, the generalized Helly property and reflexivity.

We indicate briefly some of the conventions we shall use. The sphere of radius M, $\{x \mid \| x \| \leq M\}$ is denoted by MS. The symbols x, f (or F) and x *always* refer to elements of the Banach space E and its first and second conjugate spaces, E^* and E^{**} respectively. The letters w, w^* connote the weak topology of a space as elements and as functionals respectively [2, 3, 4]. The generic $w(E)$ neighborhood of breadth ϵ of the element x_0 is $U(x_0; f_1, \cdots f_n; \epsilon) = \bigcap_{i=1}^{n} \{ x \mid | f_i(x - x_0) | < \epsilon\}$. The typical $w^*(E^*)$ neighborhood of f_0 of breadth ϵ is $U(f_0; x_1, \cdots x_n; \epsilon) = \bigcap_{i=1}^{n} \{f_0 \mid | f(x_i) - f(x_0) | < \epsilon\}$. Where a set may have several limits, these are sometimes referred to as cluster points [5]. The term *sequence* implies denumerability of a collection. A family of non null closed sets has the finite intersection property, abbreviated F. I. P., if the intersection of any finite number of the sets is non empty. The symbol $\emptyset$ is used for the null set; θ and θ^* denote the zero elements in E and E^* respectively; $\Longrightarrow$ is the symbol for implication and I is the unit interval $0 \leq t \leq 1$. We employ ( ) and [ ] to distinguish between single valued and multiple valued functions. Thus, $j(\gamma, \mu)$ is the unique value of j determined by a given γ and μ while $\mu[\gamma]$ is a typical element in a set determined by γ. $\tilde{A}$ is the complement of A. *Compact* is used in the sense of the older term bicompact [5]. Sequential $w(E)$ (conditional) compactness of A requires that any sequence $\{x^n\} \subset A$ contains a subsequence converging weakly to an element in $A(E)$. A set is compact of order $\aleph_a$ if the F. I. P. for any family $\{A_a\}$ of closed sets, of power at most $\aleph_a$, implies $\bigcap_a A_a \neq \emptyset$. For $\aleph_a = \aleph_0$ we have countable compactness. A set of the form $H_\gamma = \{x \mid f_\gamma(x) \geq d_\gamma\}$ is a *section* and A is $w(E)$ sectionally (convex) compact if the $w(A)$ closed sets

---

* Received April 10, 1941; Revised July 26, 1941; Presented to the American Mathematical Society, May 2, 1941. The principal results of this paper are stated in D. G. Bourgin, "Some properties of real linear topological spaces," *Proceedings of the National Academy of Sciences*, vol. 27 (1941), pp. 539-544.

[1] Numbers in brackets refer to the bibliography.

in the definition of compactness are restricted to the form $H_\gamma \cap A$ (convex sets) while the restriction on the cardinal numbers of the family is dropped. German capitals denote directed sets, thus $\mathfrak{A} = \{A, \succ\}$. Where the meaning is plain we often write $\{x^a\}$ for $\{x^a \mid a \in \mathfrak{A}\}$ or $\{x^a \mid \mathfrak{A}\}$. If $z^a \in E$ (E^*) is uniformly bounded for each $a \in \mathfrak{A}$ and for each $F \in E^*(E)$, $F(z^a)$ converges to a real number d_F in the Moore-Smith sense [6], then $\{z^a\}$ is a Cauchy system. For $d_F = F(z_0)$, $\{z^a\}$ converges $w(E)\,(w^*(E^*))$ to the $w(E)\,(w^*(E^*))$ limit z_0. If every Cauchy system in $A \subset E$ with $\mathfrak{A}$ arbitrary and of power $\leq \aleph_a$ is $w(E)$ convergent to an element of A, then A is $w(E)$ complete of order $\aleph_a$. Let $\{f^l \mid l \in L\}$, where L has power $\aleph_a$, be a fixed family in E^*. If for every $f \in E^*$ there is a directed set $\mathfrak{A}_f$ whose elements are members of L such that $\{f^a \mid a \in \mathfrak{A}_f\}$ is $w^*(E^*)$ convergent to f then $\{f^l\}$ is $w^*(E^*)$ c-dense of order $\aleph_a$.

2. Preliminary lemmas.

LEMMA 1. *If $A_\mu = \bigcup_{j=1}^{N_\mu} A_{\mu,j}$, where $A_{\mu,j}$ is $w(E)$, $(w^*(E^*))$ closed and N_μ is finite, and if $\{A_\mu\}$ has the F. I. P. then for each μ one can choose $A_{\mu,j(\mu)}$, a set in $\{A_{\mu,j}\}$ so that $\{A_{\mu,j(\mu)}\}$ has the F. I. P.*

This lemma is essentially a special consequence of Zorn's lemma and is originally due to Wallman [7, 8].

LEMMA 2. *The collection of all sections constitutes a sub-basis for $w(E)$ closed sets.*

If A is $w(E)$ closed,

$$\begin{aligned} A = \tilde{\tilde{A}} &= \tilde{\bigcup}_\rho U(x_\rho; f_{\rho,1}, \cdots, f_{\rho,k}; \epsilon_\rho), \qquad x_\rho \in \tilde{A} \\ &= \bigcap_\rho \bigcup_{j=1}^{2k} H_{j,\rho}. \end{aligned}$$

Here $f_{\rho,j+k} = -f_{\rho,j}$ and $H_{j,\rho} = \{x \mid f_{\rho,j}(x) - f_{\rho,j}(x_\rho) \geq \epsilon_\rho\}$, $1 \leq j \leq 2k$. Evidently $H_{j,\rho}$ is a convex, $w(E)$ closed set.

LEMMA 3. *For bounded sets the $w(E)\,(w^*(E^*))$ topology is equivalent* [9, p. 55] *to that defined by $w(E)\,(w^*(E^*))$ convergence.*

This result is obvious. We observe further that an unbounded set A may have a $w(E)\,(w^*(E^*))$ limit which is not a limit of a $w(E^*)\,(w^*(E^*))$ convergent system contained in A [4, p. 380].

3. Property (*Q*).

THEOREM 1. *Let $\{f^\mu \mid \mu \in \mathfrak{A}\}$ have θ^* as a $w^*(E^*)$ cluster point, and let A be $w(E)$ compact. Then for arbitrary $\bar{\mu} \in \mathfrak{A}$ and $\epsilon > 0$ there exists a finite set $\mu(j, \bar{\mu})$, $j = 1, \cdots N_{\bar{\mu}}$ such that $\mu(j, \bar{\mu}) \succ \bar{\mu}$ and $A = \bigcup_{j=1}^{N_{\bar{\mu}}} U(\theta; f^{\mu(j,\bar{\mu})}; \epsilon) \cap A$.*

If $x' \epsilon A$ then $U(\theta^*; x'; \epsilon)$ contains $\{f^{\mu'} \mid \mu' \epsilon \mathfrak{A}'$ cofinal in $\mathfrak{A}\}$. Hence $|f^{\mu''}(x')| < \epsilon$ for some $\mu'' \epsilon \mathfrak{A}'$ with $\mu'' \succ \mu'$. That is to say $x' \epsilon A \frown U(\theta; f^{\mu''}; \epsilon)$. Hence

$$A = \bigcup\{\mu \mid \mu \succ \bar{\mu}\}\; U(\theta; f^{\mu}; \epsilon) \frown A, \tag{1}$$

and since A is $w(E)$ compact the open covering in (1) may be replaced by a finite subcovering and this is the assertion of the theorem.

We shall say that A has property (Q) if, whenever $\{f^{\gamma,\mu} \mid \mu \epsilon \mathfrak{A}_\gamma\}$ is $w^*(E^*)$ *convergent* to f_γ, then for arbitrary $\bar{\mu} \epsilon \mathfrak{A}_\gamma$ and $\epsilon > 0$ there exists a finite set $\{\mu(j, \bar{\mu}) \mid \mu(j, \bar{\mu}) \succ \bar{\mu},\ j = 1, \cdots, N_{\gamma,\bar{\mu}}\}$ such that $A = \bigcup_{j=1}^{N_{\gamma,\bar{\mu}}} U(\theta; f^{\gamma,\mu(j,\bar{\mu})} - f_\gamma; \epsilon) \frown A$. If the property (Q) holds for $w^*(E^*)$ convergent systems of power $\leqq \aleph_a$ we write property Q_a. Evidently Theorem 1 may be extended to assert that $w(E)$ compactness of order $\aleph_a$ implies property Q_a.

In the expression for property (Q) assume that ϵ is replaced by $\epsilon_{\bar{\mu}}$, A by S and $\epsilon_{\bar{\mu}} \to 0$ in the sense of Moore-Smith convergence. We define $B_{\gamma,j,\bar{\mu}}$ by $B_{\gamma,j,\bar{\mu}} = S \frown \{x \mid f^{\gamma,\mu(j,\bar{\mu})}(x) \geqq f_\gamma(x_\gamma) + \epsilon_\gamma - \epsilon_{\bar{\mu}}\}$. The following theorem is used in the proof of Theorem 22.

THEOREM 2. *If S has property (Q) then $\bigcap_{\bar{\mu}} \bigcup_{j=1}^{N_{\gamma,\bar{\mu}}} B_{\gamma,j,\bar{\mu}} = K_\gamma$, where $K_\gamma = S \frown \{x \mid f_\gamma(x) \geqq f_\gamma(x_\gamma) + \epsilon_\gamma\}$.*

Let $x' \epsilon K_\gamma$. For fixed $\bar{\mu}$, property (Q) implies that for some j, denoted by J, $1 \leqq J \leqq N_{\gamma,\bar{\mu}}$,

$$|f_\gamma(x') - f^{\gamma,\mu(J,\bar{\mu})}(x')| < \epsilon_{\bar{\mu}}, \tag{2}$$

that is to say either $f^{\gamma,\mu(J,\bar{\mu})}(x') \geqq f_\gamma(x')$ or $\epsilon_{\bar{\mu}} \geqq f_\gamma(x') - f^{\gamma,\mu(J,\bar{\mu})}(x') \geqq 0$. On making use of the definition of K_γ it follows that $f^{\gamma,\mu(J,\bar{\mu})}(x') \geqq f_\gamma(x_\gamma) + \epsilon_\gamma - \epsilon_{\bar{\mu}}$. Hence

$$\bigcup_{j=1}^{N_{\gamma,\bar{\mu}}} B_{\gamma,j,\bar{\mu}} \supset K_\gamma. \tag{2.1}$$

Denote the left side of the theorem's asserted equality by P_γ. That $P_\gamma \supset K_\gamma$ follows from (2.1). Suppose that $x_0 \epsilon P_\gamma$; then, by the definition of $B_{\gamma,j,\bar{\mu}}$, for each $\bar{\mu}$ there is a value of j, denoted by $J(\bar{\mu}, \gamma)$, $1 \leqq J(\bar{\mu}, \gamma) \leqq N_{\gamma,\bar{\mu}}$ such that

$$f^{\gamma,\mu(J(\bar{\mu},\gamma),\bar{\mu})}(x_0) \geqq f_\gamma(x_\gamma) + \epsilon_\gamma - \epsilon_{\bar{\mu}}. \tag{2.2}$$

For arbitrary δ, $\bar{\mu}_1$ exists such that $|f^{\gamma,\mu}(x_0) - f_\gamma(x_0)| < \epsilon_{\bar{\mu}} < \delta$ for all $\mu \succ \bar{\mu}_1$ and $\bar{\mu} \succ \bar{\mu}_1$. Since $\mu(J(\bar{\mu}, \gamma), \bar{\mu}) \succ \bar{\mu}$ (Cf. (Q)) the set $\{\mu(J(\bar{\mu}, \gamma), \bar{\mu})\}$

is cofinal in $\mathfrak{A}_\gamma$ and accordingly $| f^{\gamma,\mu(J(\bar{\mu},\gamma),\bar{\mu})}(x_0) - f_\gamma(x_0) | < \epsilon_{\bar{\mu}} < \delta$ for all $\mu \succ \mu_1$. Hence (2.2) implies $f_\gamma(x_0) \geqq f_\gamma(x_\gamma) + \epsilon_\gamma$ or $x_0 \in K_\gamma$ and $P_\gamma \subset K_\gamma$.

THEOREM 3. *Let $C(B)$ be the Banach space of continuous functions $x(b \mid B)$ on the compact Hausdorff space B. Let V be a bounded subset of $C(B)$ with property (Q). If $\{x^a \mid \mathfrak{A}\} \subset V$ converges pointwise for the set T dense in B then it converges pointwise (on B) to $x_0(b) \in C(B)$.*

Let $y(b \mid T)$ denote the limit function of $\{x^a \mid \mathfrak{A}\}$ on T. We must show (a) that $y(b \mid T)$ may be extended to the continuous function $y(b \mid B)$ and (b) that $\{x^a(b \mid B) \mid \mathfrak{A}\}$ converges pointwise to $y(b \mid B) = x_0(b \mid B)$. Suppose $b_0 \in B - T$ and let $\{b_\gamma \mid b_\gamma \in T, \gamma \in \mathfrak{G}\}$ converge in the Moore-Smith sense to b_0. If (a) is untrue, for some $\epsilon > 0$, there are sets $\mathfrak{G}'$ and $\mathfrak{G}''$ cofinal in $\mathfrak{G}$ such that for each $\gamma' \in \mathfrak{G}'$ there is a $\gamma'' = \gamma''(\gamma') \in \mathfrak{G}''$ with $\gamma'' \succ \gamma'$ for which,

$$(3) \qquad | y(b_{\gamma''}) - y(b_{\gamma'}) | > \epsilon$$

Now

$$(3.1) \qquad | y(b_{\gamma''}) - y(b_{\gamma'}) | \leqq | y(b_{\gamma''}) - x^a(b_{\gamma''}) | + | x^a(b_{\gamma''}) - x^a(b_{\gamma'}) | + | x^a(b_{\gamma'}) - y(b_{\gamma'}) | .$$

Evidently $x(b_{\gamma'}) - x(b_{\gamma''}) = f^{\gamma'}(x)$ is a linear functional on $C(B)$. Since $(| x(b_{\gamma''}) - x(b_{\gamma'}) | / \sup_{b \in B} | x(b) |) \leqq 2$, it is clear that $\{f^{\gamma'}\}$ is uniformly bounded and is $w^*(C^*(B))$ convergent to θ^*. Property (Q) implies that there is a cofinal set $\mathfrak{G}'''$ in $\mathfrak{G}'$ such that $| f^{\gamma'''}(x^a) | < \epsilon/3$ for a set $\mathfrak{A}_{\gamma'''}$ (associated with each $\gamma''' \in \mathfrak{G}'''$) cofinal in $\mathfrak{A}$. By hypothesis the first and last absolute values on the right side of (3.1) are dominated by $\epsilon/3$ for $a \succ a_0$. Thus for $\{a \mid a \in \mathfrak{A}_{\gamma'''}, a \succ a_0\}$ the left side of (3.1) is inferior to ϵ in contradiction with (3). Hence $y(b_\gamma)$ converges in the Moore-Smith sense and we denote the limit by $y(b_0)$. That is to say the extended function $y(b \mid B)$ is continuous.

Assume that (b) is false. Then for some set $\mathfrak{A}'$ cofinal in $\mathfrak{A}$

$$(3.2) \qquad | x^a(b_0) - y(b_0) | > \epsilon \quad \text{for all } a \in \mathfrak{A}'.$$

Now

$$(3.3) \qquad | x^a(b_0) - y(b_0) | \leqq | x^a(b_0) - x^a(b_\gamma) | + | x^a(b_\gamma) - y(b_\gamma) | + | y(b_\gamma) - y(b_0) | .$$

By the continuity of $y(b \mid B)$ we have $| y(b_\gamma) - y(b_0) | < \epsilon/3$ for $\gamma \succ \gamma_0$. Now $F^\gamma(x) = x(b_\gamma) - x(b_0)$ is a linear functional and $\{F^\gamma\}$ converges

$w^*(C^*(B))$ to θ^*. Thus $|x^a(b_0) - x^a(b_\gamma)| < \epsilon/3$ for $\gamma \in \mathfrak{G}_1$ cofinal in $\mathfrak{G}$ and $a \in \mathfrak{A}_\gamma$ cofinal in $\mathfrak{A}'$. Let $\gamma_1 \in \mathfrak{G}_1$ with $\gamma_1 \succ \gamma_0$ and choose $\bar{a} \in A_{\gamma_1}$ so that $|x^a(b_{\gamma_1}) - y(b_{\gamma_1})| < \epsilon/3$, then for $a = \bar{a}$ the left side of (3.3) is dominated by ϵ in contradiction with (3.2). (The boundedness condition on V is superfluous in view of Theorem 8).

The following two theorems are stated without proof by Sirvint [10].

THEOREM 4. *$A \subset C(I)$ is $w(E)$ sequentially, conditionally compact if and only if A has property (Q_0).*

THEOREM 5. *For separable E, $A \subset E$ is $w(E)$ sequentially, conditionally compact if and only if A has property (Q_0).*

Since we make use of these theorems it seems worth while to give proofs. Suppose that either necessity assertion is false. Then for some sequence $\{f^n\}$ converging $w^*(E^*)$ to θ^* and some ϵ and $\bar{n}$ we can find $x^n \in A$ to satisfy $x^n \bar{\epsilon} \bigcup_{i=\bar{n}}^{n} U(\theta; f^i; \epsilon)$. Let x_0 be a $w(E)$ cluster point of $\{x^n\}$. Manifestly we can select f^m, $m > \bar{n}$, from $\{f^n\}$ and x^p, $p > m$, from $\{x^n\}$ to satisfy $|f^m(x_0)| < \epsilon/2$ and $|f^m(x^p - x_0)| < \epsilon/2$. Hence $|f^m(x^p)| < \epsilon$ in contradiction with the definition of x^p. For the sufficiency argument in Theorem 4, we take T in Theorem 3 as the set of rational points $\in I$. If $\{x^n\} \subset A$, an obvious application of the Cantor diagonal process yields a subsequence $\{x^r(t)\}$ converging pointwise on T. By Theorem 3 $\{x^r(t)\}$ converges pointwise to $y(t) \in C(I)$. Since $w(E)$ sequential convergence is the same as pointwise convergence in $C(I)$ [1, p. 134], the proof of Theorem 4 is complete.

Concerning Theorem 5: E is equivalent to $E_1 \subset C(I)$ [1, p. 185] under the linear operator V. To every $F \in C^*(I)$ there corresponds an $f \in E^*$ for which

$$f(x) = F(y), \tag{5}$$

where $y = V(x)$ and x, y belong to E and $C(I)$ respectively [1, p. 100]. Hence $w^*(C^*(I))$ convergence of $\{F^n\}$ implies $w^*(E^*)$ convergence of the corresponding $\{f^n\}$. Hence property (Q_0) for A implies property (Q_0) for $A_1 = V(A)$ and therefore by Theorem 4, A_1 is $w(C(I))$ sequentially compact. Let $\{x^n\} \subset A$ and $\{y^n\} \subset A_1$ correspond. We know that for some subsequence, $F(y^{n_i}) \to F(y_0)$, $y_0 \in V(E)$ [1; p. 57], for all F and therefore by (5) and [1; p. 188] $f(x^{n_i}) \to f(x_0)$, $x_0 \in E$, for all f. Hence A is $w(E)$ sequentially, conditionally compact.

4. $w(E)$ compactness.

THEOREM 6. (a) *A is $w(E)$ compact $\Longleftrightarrow$ A is $w(E)$ convex compact.*[2]

(b) *A is $w(E)$ compact $\Longleftrightarrow$ A is $w(E)$ sectionally compact.*

The direct implications are trivial. The inverse implication in (b) includes that in (a) and follows at once from Lemmas 1 and 2.

THEOREM 7. *If K is bounded, convex and strongly closed, then K is $w(E)$ compact $\Longleftrightarrow$ K is transfinitely closed.*

The proof is a straightforward consequence of Theorem 6 part (a) and a result of Milman's [11].

THEOREM 8. *If A, $A \subset E$, is $w(E)$ compact, convex compact or sectionally compact of order $\aleph_a$ or has property Q_a, then A is bounded.*

If A were not bounded then we could find a sequence $\{x^n\}$ with $\| x^n \| = M_n \uparrow \infty$. For some f, say f_0, lim sup $f_0(x^n) = \infty$ [1; p. 80]. This is easily seen to be inconsistent with property Q_0 when the sequence $\{f_0/N \mid N = 1, 2, \cdots\} \to \theta^*$ is used.

THEOREM 9. *If $A \subset E$ is $w(E)$ compact of order $\aleph_a$, then A is $w(E)$ complete of order $\aleph_a$.*

Let $\{x^\mu \mid \mu \in \mathfrak{B}\}$ be an arbitrary Cauchy system in A, with $\mathfrak{B}$ of power $\leq \aleph_a$. Let $F_{\bar{\mu}} = \{x^\mu \mid \mu \succ \bar{\mu}\}$ and write $\bar{F}_{\bar{\mu}}$ for the $w(A)$ closure. Then since $\{\bar{F}_{\bar{\mu}}\}$ has the F. I. P. and A is $w(E)$ compact of order $\aleph_a$, there is at least one common point, say x_0. By the definition of a Cauchy system $f(x^\mu) \to d_f$. Hence for each $\epsilon > 0$ $\exists$ μ_ϵ such that $| f(x^{\mu'}) - f(x^{\mu''}) | < \epsilon$ for all successors, μ' and μ'', to μ_ϵ. Let $x^{\bar{\mu}} \in U(x_0; f; \epsilon)$, $\bar{\mu} \succ \mu_\epsilon$. Then clearly $\{x^{\mu_1}\} \in U(x_0; f; 2\epsilon)$ where $\mu_1 \succ \bar{\mu}$. Thus $\{\mu_1\}$ is a residual set [5, p. 10] in $\mathfrak{B}$. Manifestly, then, every $w(E)$ neighborhood of x_0 contains a subset of $\{x^\mu\}$ corresponding to a residual μ set in $\mathfrak{B}$. This is precisely the condition for $w(E)$ convergence to x_0

---

[2] Theorem 6a is apparently known (at least under the superfluous restrictions on A of boundedness and closure) and so is Theorem 7. However they are not given in the literature.

(and the fact that E is a Hausdorff space in its $w(E)$ topology implies that x_0 is a unique limit).

5. Iterated limits.

THEOREM 10. *If* (a) $\{f^{\mu}\}$ *is total* [1, p. 42] *and* (b) $\{x^{\rho}\}$, $\rho \epsilon \mathfrak{W}$ *a directed set of power* $\leq \aleph_a$, *is contained in a* $w(E)$ *compact* (*of order* $\aleph_a$) *set, and* (c) $f^{\mu}(x^{\rho})$ *converges in the Moore-Smith sense as a set of real numbers to* $f^{\mu}(x_0)$ *for each* μ, *then* $\{x^{\rho}\}$ *is* $w(E)$ *convergent to* x_0.

Let $A_{\bar{\rho}}$ be the $w(A)$ closure of $\{x^{\rho} \mid \rho \succ \bar{\rho}\}$. Condition (b) implies common elements for $\{A_{\bar{\rho}}\}$. Suppose that one of these common elements, say $\bar{x}$, differs from x_0. Write $\mathfrak{W}_{\epsilon}$ for $(\{\rho \mid x^{\rho} \epsilon U(\bar{x}; f^{\mu}; \epsilon\}, \succ)$ where we assume the ordering in $\mathfrak{W}_{\epsilon}$ to be consistent with that in $\mathfrak{W}$. Since $\bar{x}$ is a $w(E)$ cluster point $\mathfrak{W}_{\epsilon}$ is cofinal in $\mathfrak{W}$ for every $\epsilon > 0$. This is possible only if $f^{\mu}(\bar{x} - x_0) = 0$. However $\{f^{\mu}\}$ is *total* wherefore $\bar{x} = x_0$.

In view of Theorem 8, $\{x^{\rho}\}$ is uniformly bounded. We have yet to show that $\{x^{\rho}\}$ is $w(E)$ convergent to x_0. If this were not so, then for some $\epsilon > 0$ there would be a neighborhood $U(x_0; f_1, \cdots, f_n; \epsilon)$ (f_i need not be in $\{f^{\mu}\}$) such that $\mathfrak{W}_1 = (\{\rho \mid x^{\rho} \bar{\epsilon} U(x_0; f_1, \cdots, f_n; \epsilon)\}, \succ)$, under the ordering imposed by $\mathfrak{W}$, is cofinal in $\mathfrak{W}$. As a consequence of (b), $\{x^{\rho} \mid \rho \epsilon \mathfrak{W}_1\}$ has a $w(E)$ cluster point, say x'. It then follows, just as for $\bar{x}$, that $x_0 = x'$. That is to say, for some $\rho \epsilon \mathfrak{W}_1$ we have $| f_i(x^{\rho}) - f_i(x_0) | < \epsilon/2$, $i = 1, \cdots, n$, in contradiction with the definition of $\mathfrak{W}_1$. Hence $\mathfrak{W}_1$ cannot be cofinal in $\mathfrak{W}$ and the theorem is established.

COROLLARY. *If in Theorem* 10, *condition* (a) *is changed to* (a′) $\{f^{\nu}\} \subset \{f^{\mu}\}$ *converges* $w^*(E^*)$ *to* f_0 *and* $\{f^{\mu}\}$ *is total, the other conditions remaining the same, then* $L_{\rho}L_{\nu}f^{\nu}(x^{\rho}) = L_{\nu}L_{\rho}f^{\nu}(x^{\rho}) = f_0(x_0)$, *where the limits are taken in the Moore-Smith sense.*

COROLLARY. *If* A *is a* $w(E)$ *sequentially compact subset of the separable Banach space* E *and if* $f^j(x^m) \rightarrow f^j(x_0)$, $j = 1, \cdots$ *where* $\{f^j\}$ *is total and* $\{x^m\} \subset A$, *then* $\{x^m\}$ *is* $w(E)$ *sequentially convergent to* x_0 *and* $x \epsilon A$.[8]

THEOREM 11. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *have elements* a *and* b *respectively. If* A, *contained in* E, *has property* (*Q*) *and if* (a) $\{x^a \mid \mathfrak{A}\} \subset A$ *converges* $w^*(E^{**})$

---

[8] The condition $\aleph_a = \aleph_0$ in Theorem 10 implies that A is $w(E)$ countably compact. However, by Theorem 13, this is the same as $w(E)$ sequentially compact.

to X_0 *and* (b) $\{f^b \mid \mathfrak{B}\}$ *converges* $w^*(E^*)$ *to* f_0 *then* $L_bL_af^b(x^a) = L_aL_bf^b(x^a) = X_0(f_0)$, *where the limits are taken in the Moore-Smith sense.*[4]

Suppose that the conclusion is untrue. Then

$$(11) \qquad | X_0(f^b) - X_0(f_0) | > \epsilon, \quad b \epsilon \mathfrak{B}' \qquad \text{cofinal in } \mathfrak{B}.$$

In view of property (Q) there exist $b' \epsilon \mathfrak{B}'$ and a set $\mathfrak{A}_{b'}$ cofinal in $\mathfrak{A}$ such that $| f^{b'}(x^a) - f_0(x^a) | < \epsilon/3$ for all $a \epsilon \mathfrak{A}_{b'}$. For $a > a_0$, $| f_0(x^a) - X_0(f_0) | < \epsilon/3$ by reason of (a). Again because of (a), $| X_0(f^{b'}) - f^{b'}(x^{a'}) | < \epsilon/3$ for some $a' \epsilon \mathfrak{A}_{b'}$ with $a' > a_0$. On making use of the triangle inequality we derive a contradiction with (11) for $b = b'$.

6. $w(E)$ sequential compactness.

THEOREM 12. *If* E *is separable,* $A \subset E$ *is* $w(E)$ *sequentially compact and* $\Gamma = \{F^n\} \subset E^*$ *is* $w^*(E^*)$ *dense in* E^* [1, p. 124] *then,* (a) $\{V_n(x_0)\}$, *where* $V_n(x_0) = A \cap U(x_0; F', \cdots, F^n; 1/n)$, *constitutes a* $w(E)$ *neighborhood basis of* x_0 *and* (b) A *satisfies the second denumerability axiom.*

Suppose that (a) is untrue. Then there is a neighborhood $U' = U'(x_0, f_1, \cdots, f_n; \epsilon) \cap A$ not containing any neighborhood of the asserted basis. Hence for each $n \ni x^n \epsilon V_n(x_0) \cap \bar{U}'$. A subsequence $\{x^{n_i}\}$ converges $w(E)$ to $\bar{x}$ since A is $w(E)$ sequentially compact. Moreover $| F^j(x^{n_i} - x_0) | < 1/n_i$; for $j \leq n_i$. Hence $F^j(x^{n_i}) \to F^j(x_0)$ for each j and by the second corollary to Theorem 10 $\bar{x} = x_0$. Accordingly U' contains elements in $\{x^{n_i}\}$ in contradiction with the definition of $\{x^n\}$.

Let the denumerable set $\{x^r\}$ be strongly dense in A. Consider $V_n(x_0)$. Let $x^p \epsilon \{x^r\}$ satisfy $|| x^p - x_0 || < 1/2n \max (|| F^i ||)$, $i = 1, \cdots, 2n$. Then it is easy to show that $x_0 \epsilon V_{2n}(x^p) \subset V_n(x_0)$. Accordingly the collection of the bases for all elements of $\{x^r\}$ constitutes a basis for all open sets in A and is plainly denumerable.

THEOREM 13. *For subsets of a separable Banach space* E, *the concepts of* $w(E)$ *sequential compactness and* $w(E)$ *compactness and countable compactness are equivalent.*

Suppose, then, that $A \subset E$ is $w(E)$ sequentially compact and let $\{G^n\}$ be a denumerable open covering. If $x^n \bar{\epsilon} \bigcup_{i=1}^{n} G^i$ then there is a $w(E)$ cluster point x_0. If $x_0 \epsilon G^m$ then some $w(E)$ neighborhood of x_0 is included in G^m and contains an infinite subsequence of $\{x^n\}$, that is to say G^m contains

---

[4] Condition (a) is to be understood in the sense that if X_a is defined by $X_a(f) = f(x_a)$ for all $f \epsilon E^*$, then $\{X_a\}$ converges $w^*(E^{**})$ to X_0.

$x^{n'}$ where $n' > m$ which contradicts the definition of $\{x^n\}$. On the other hand if A is $w(E)$ countably compact, it has property (Q_0), and hence, in view of Theorem 5, A is $w(E)$ sequentially conditionally compact. In view of Theorem 12(b) it now follows that $w(E)$ countably compact sets are $w(E)$ closed. Thus A is actually $w(E)$ sequentially compact.

Theorem 14. *If A, $A \subset E$, is $w(E)$ compact of order $\aleph_a$ then A is $w(E)$ sequentially compact, while if A is $w(E)$ sequentially compact, then A is $w(E)$ countably compact.*

Let $\{x^n\}$ be chosen arbitrarily in A and let E_0 be the smallest closed linear manifold spanning $\{x^n\}$. Write A_0 for $A \cap E_0$. If $\{A_0{}^m\} \subset A_0$ is a sequence of $w(A_0)$ closed sets with the F. I. P. then since E_0 is necessarily $w(E)$ closed [1, p. 57] it follows that $A_0{}^m$ is $w(A)$ closed and $\bigcap_m A_0{}^m \neq \emptyset$. Therefore A_0 is $w(E_0)$ countably compact. By Theorem 13, A_0 *is* $w(E_0)$ sequentially compact. Hence $\{x^{n_i}\} \subset \{x^n\}$ converges $w(E_0)$ to $x_0 \in A_0$. Each $f \in E^*$ determines $f' \in E_0{}^*$ such that $f(x) = f'(x)$ for all $x \in E_0$. Hence $L_{i \to \infty} f(x^{n_i}) = f(x_0)$ and therefore $\{x^{n_i}\}$ is $w(E)$ convergent to $x_0 \in A_0 \subset A$.

The second half of the theorem is covered by the demonstration given for Theorem 13.

7. **The extended Helly property.** The Helly theorem [12, 3′] provides a powerful tool for treating many of the problems of separable Banach spaces in a unified manner. It is therefore natural to attempt to remove the countability restriction. We show that the validity of a natural extension of the Helly theorem is closely related to the compactness and completeness characterizations central in this paper. By the extended Helly property of order $\aleph_a$ for E we shall mean that under the condition (H), $|\sum_{i=1}^N \lambda_{\mu_i} c_{\mu_i}| \leq M \| \sum_{i=1}^N \lambda_{\mu_i} f^{\mu_i} \|$, where $\{\lambda_{\mu_i} \mid i = 1, \cdots N\}$ is *any* finite set of real numbers, the system $f^\mu(x) = c_\mu$ consisting of at most $\aleph_a$ equations has a solution $x_\epsilon \in (M + \epsilon) S$ for each $\epsilon > 0$.

Theorem 15. *If S is $w(E)$ compact of order $\aleph_a$ then E has the extended Helly property of order $\aleph_a$.*

For fixed $\epsilon > 0$ let A_μ be the subset of $(M + \epsilon) S$ consistent with the single equation $f^\mu(x) = c^\mu$. Plainly A_μ is convex and strongly closed and is therefore $w(E)$ closed [13]. Moreover, $\{A_\mu\}$ has the F. I. P. for Helly's theorem guarantees solutions, interior to $(M + \epsilon) S$, $\epsilon > 0$, of any finite set of equations $f^{\mu_i}(x) = c^{\mu_i}$, $i = 1, \cdots N$, when (H) is satisfied. Since S, and therefore $(M + \epsilon) S$ also, is $w(E)$ compact of order $\aleph_a$, it follows that $\bigcap_\mu A_\mu \neq \emptyset$ which is tantamount to the conclusion sought.

THEOREM 16. *If* E *has the extended Helly property for all orders then* E *is* $w(E)$ *complete.*

Let $\{x^a \mid \mathfrak{A}\}$ be a Cauchy system which is uniformly bounded by M. Let $\{f^\mu\}$ be $w^*(E^*)$ c-dense in E^*. Consider the system,

$$\{f^\mu(x) = d_\mu\}, \quad d_\mu = \lim f_\mu(x^a). \tag{16}$$

Now

$$\left| \sum_{i=1}^N \lambda_{\mu_i} d_{\mu_i} \right| \leq \left| \sum_{i=1}^N \lambda_{\mu_i} (d_{\mu_i} - f^{\mu_i}(x^a) \right| + \left| \sum_{i=1}^N \lambda_{\mu_i} f^{\mu_i}(x^a) \right|. \tag{16.1}$$

For arbitrary $\delta > 0$ and prescribed λ_{μ_i}, $i = 1, \cdots, N$, we can choose $\epsilon > 0$ to satisfy $\epsilon N \max \{\lambda_{\mu_i}\} < \delta$. In view of the convergence of $f^{\mu_i}(x^a)$ to d_{μ_i} for $i = 1, 2, \cdots, N$, we can find $a_0 \epsilon \mathfrak{A}$ for which $\left| \sum_{i=1}^N \lambda_{\mu_i} (d_{\mu_i} - f^{\mu_i}(x^{a_0}) \right| < \delta$. Hence the right side of (16.1) is dominated by

$$\delta + \sup_{x^* \epsilon MS} \left| \sum_{i=1}^N \lambda_{\mu_i} f^{\mu_i}(x^a) \right| \leq \delta + M \left\| \sum_{i=1}^N \lambda_{\mu_i} f^{\mu_i} \right\|,$$

and since $\delta > 0$ is arbitrary, condition (H) is satisfied by the system (16). Accordingly a solution $x_\epsilon \epsilon (M + \epsilon) S$ exists, for each $\epsilon > 0$. Moreover since $\{f^\mu\}$ is obviously total, x_ϵ is independent of ϵ and its uniqueness is connoted by writing $x_\epsilon = x_0$. A familiar argument now guarantees that $f(x_0) = \lim f(x^a)$ for arbitrary f. Indeed our previous reasoning shows that the augmented system consisting of (16) and $f(x) = d_f$ must have a solution which, plainly, is again x_0. This concludes the proof.

COROLLARY. *If* E *has the extended Helly property of order* $\aleph_a$ *and* E^* *contains a total set of power* $\aleph_a$ *then* S *is* $w(E)$ *complete.*

The proof given above for Theorem 16 obviously carries over.

8. Sets in E^*.

THEOREM 17. *If* $\Gamma \subset E^*$ *is strongly compact then* Γ *is* $w^*(E^*)$ *compact.*[5]

Since metric compactness implies boundedness and bounded, $w^*(E^*)$ closed sets are $w^*(E^*)$ compact [2, 3′] we need merely show that Γ is $w^*(E^*)$ closed. Suppose that $f_0 \epsilon \bar{\Gamma}$ where $\bar{\Gamma}$ is the $w^*(E^*)$ closure of Γ. Let $\boldsymbol{U} = \{U^\mu\}$ be a $w^*(E^*)$ neighborhood basis at f_0 where $\mu \epsilon \mathfrak{A}$ and the ordering in $\mathfrak{A}$ is given by $\mu \succ \mu' \Longleftrightarrow U^\mu \subset U^{\mu'}$. For each U^μ let f^μ be any element in $U^\mu \frown \Gamma$. Then $\{f^\mu\}$ is $w^*(E^*)$ convergent to f_0. In view of the total boundedness of Γ there is a closed sphere σ_1 in Γ of diameter ϵ_1 containing $\{f^{\mu'} \mid \mu' \epsilon \mathfrak{A}'$ cofinal in $\mathfrak{A}\}$. Similarly $\sigma_2 \subset \sigma_1$, σ_2 of diameter $\epsilon/2$ contains $\{f^{\mu^2} \mid \mu^2 \epsilon \mathfrak{A}^2$ cofinal in $\mathfrak{A}'\}$, etc. A unique element $\bar{f} \epsilon \Gamma$ is defined by $\bigcap_n \sigma_n$. For arbitrary δ

---

[5] This theorem was suggested by somewhat similar results obtained by Kakutani for the space E.

and any finite collection $\{x_i\}$ $i = 1, \cdots, n$, $\delta \geq \| f^\mu - \bar{f} \| \| x_i \| \geq | f^\mu(x_i) - \bar{f}(x_i) |$ for all $\mu \in \mathfrak{A}m(\delta, \{x_i\})$. Hence any $w^*(E^*)$ neighborhood of $\bar{f}$ contains a sphere $\sigma_m \in \{\sigma_n\}$ and this implies that $\bar{f}$ is a $w^*(E^*)$ cluster point of $\{f^{\mu'}\}$ and hence of $\{f^\mu \,|\, \mathfrak{A}\}$ also. Since $w^*(E^*)$ limits of convergent systems in the Hausdorff space E^* are necessarily unique, it follows that $\bar{f} = f_0$.

The following two theorems are generalizations of Banach's Theorem 2, p. 213.

THEOREM 18. *Let Γ be a linear subspace of E^*. A necessary and sufficient condition for $\bar{\Gamma} = E^*$, where $\bar{\Gamma}$ is the $w^*(E^*)$ closure of Γ, is that for arbitrary x, $f(x) = \| x \|$ for some $f \in \Gamma$.*[6]

Assuming $x_0 \neq \theta$, if $f(x_0) = c = 0$ then $(\| x_0 \| / c) f = f' \in \Gamma$ and $f'(x_0) = \| x_0 \|$. If $f(x_0) = 0$ for all $f \in \Gamma$ let $f_0 \in \bar{\Gamma}$ with $f_0(x_0) = \| x_0 \|$ [1; p. 55]. However $| f_0(x_0) - f(x_0) | < \epsilon$ for some $f \in \Gamma$, since $\bar{\Gamma} = E^*$, and accordingly we have the contradictory implication $f_0(x_0) = 0$. On the other hand assume that $\bar{\Gamma} \neq E^*$ under the condition stated in the theorem. Since $\bar{\Gamma}$ is a $w^*(E^*)$ closed linear subspace it is also regularly closed [3]. Hence if $f_0 \bar{\in} \bar{\Gamma}$ there exists an x_0 for which

$$f_0(x_0) = 1 \tag{18}$$

$$f(x_0) = 0, \quad f \in \bar{\Gamma}. \tag{18.1}$$

Since $\| x_0 \| \neq 0$, by (18), it is clear that (18.1) contradicts the assumption that $f(x_0) = \| x_0 \|$ for some $f \in \Gamma$.

THEOREM 19. *Let Γ be a linear subspace of E^* and denote by Γ' its $w^*(E^*)$ convergence closure. A necessary and sufficient condition for $\Gamma' = E^*$ is that there exist a fixed sphere $\Gamma_M = MS^* \cap \Gamma$ such that for each x and some $f \in \Gamma_M$, $f(x) = \| x \|$.*

To establish sufficiency assume the conclusion untrue. Let f_0 be absent from Γ'. Then $F_0 = (f_0 / \|f_0\|) \bar{\in} \Gamma'_M$. For bounded sets the $w^*(E^*)$ and the $w^*(E^*)$ convergence closures are identical by Lemma 3. Hence Γ'_M is $w^*(E^*)$ closed and is obviously convex. This implies that Γ'_M is *regularly convex* [14]. Accordingly for some x_1

---

[6] The second condition in Banach's theorem, namely that $f(x) = || x ||$ holds for f restricted to $MS^* \cap \Gamma$ for some fixed M, is unavailable here. Indeed if we were to attempt to establish such a condition by the analogue of Banach's demonstration of necessity, we should have, with $\bar{\Gamma}_n = \Gamma \cap nS^*$, $\overline{\bigcup_n \bar{\Gamma}_n} \subset \bigcup \Gamma_n = E^*$ and the inclusion may be proper (Cf. the comment on Lemma 3). Hence to preserve complete analogy with Banach's statement a closure operation involving bounded sets alone, is required (Cf. Theorem 19).

$$\| x_1 \| = \| F_0 \| \, \| x_1 \| > \sup_{f \epsilon \Gamma_{M'}} f(x_1) \geq \sup_{f \epsilon \Gamma_M} f(x_1).$$

This contradicts the hypothesis of the theorem.

The necessity proof may be obtained from the argument of Banach, pp. 213-215, by formally replacing $w^*(E^*)$ sequential convergence by $w^*(E^*)$ convergence.

9. **An example.** Sequences ultimately of the form $(x, x, x, \cdots)$ are termed *trivial.* There is a classic example, due to von Neumann [4, p. 380] of a sequence in l_2 with a $w(E)$ limit point which is not a $w(E)$ sequential limit point. This example pivots on the non-boundedness of the sequence. We propose below to construct a *bounded* $w^*(E^*)$ set which is even *compact* and yet contains no non-trivial $w^*(E^*)$ convergent sequences.

We first summarize in slightly different guise some known results [15, 5]. Let $T_i \subset I$, $i = 1, \cdots$ denote the set $\bigcup_{k=1}^{2^{i-1}} \{t \mid (2k-1)/2^i \leq t < k/2^{i-1}\} \cup \{1\}$ and write $g^i(t)$ for the characteristic function of T_i. Let $G = \{g^i(t)\}$. Then G is a sub-space of the Boolean product space $\Pi\{\psi_t \mid t \epsilon I, \psi_t = 0 \text{ or } 1\}$ denoted by $2^{|I|}$, which is the space of all characteristic functions $b(t)$ on I. The generic neighborhood of $b_0(t) \epsilon 2^{|I|}$ is $\{b(t) \mid b(t_i) = b_0(t_i), i = 1, \cdots N\}$; Moore-Smith convergence of $\{b^\mu(t) \mid \mu \epsilon \mathfrak{A}\} \Longleftrightarrow$ Moore-Smith pointwise convergence of $b^\mu(t)$; and $2^{|I|}$ is a compact Hausdorff space. Hence $\bar{G}$, the closure of G in $2^{|I|}$, is a compact Hausdorff space in which sequential convergence is equivalent to pointwise convergence. The elements of $\bar{G}$ are denoted by $g(t)$.[7] Finally $\bar{G}$ contains no convergent sequences [15].

Let $C(\bar{G})$ be the space of continuous functions $x(g)$ on $\bar{G}$ with the usual norm, namely $\| x \| = \max_{g \epsilon \bar{G}} | x(g) |$. Define $f_g \epsilon C^*(\bar{G})$ by $f_g(x) = x(g)$, Cf. [3′]. Then $\| f_g \| = \sup_{\|x\| \leq 1} | f_g(x) | = \sup_{\|x\| \leq 1} | x(g) | \leq 1$. We assert that $\{f_g\}$ is a $w^*(E^*)$ compact subset of S^* homeomorphic to $\bar{G}$. Evidently $x(g) = g(t_0) \epsilon C(\bar{G})$. Now $\{f_g\}$ and $\bar{G}$ are in one to one correspondence for $f_{g'} \neq f_{g''} \Longrightarrow g' \neq g''$ while if $g' \neq g''$ we may choose $t_0 \epsilon I$ so that $g'(t_0) \neq g''(t_0)$ and hence, if $x(g) = g(t_0)$, $x(g') \neq x(g'')$ or finally $f_{g'} \neq f_{g''}$. We indicate the correspondence between $\bar{G}$ and $\{f_g\}$ by $f_g = \psi(g)$. Since $\bar{G}$ is compact and ψ is easily seen to be continuous it is clear that ψ defines a homeomorphism. Hence $\{f_g\}$ is $w^*(E^*)$ compact.

---

[7] The writer conjectures, incidentally, that the effective definition of an element of $\bar{G} \cap \bar{G}$ (in Sierpinski's sense) is tantamount to the effective construction of a non-measurable (Lebesgue) set. Since this was written Professors Kakutani and Eilenberg have communicated a proof to the writer which shows that such elements are characteristic functions of non-measurable sets.

Suppose that $\{f_{g_n}\}$ is $w^*(E^*)$ convergent to $\{f_{g_0}\}$ for a non-trivial sequence $\{g_n\}$. This implies that $x(g_n) \rightarrow x(g_0)$ for all $x \in C(\bar{G})$, and in particular for each x of the type $x(g) = g(t)$ for an arbitrary fixed $t \in I$. Thus $\{g_n(t)\}$ would be pointwise convergent to $g_0(t)$ which is impossible since $\bar{G}$ contains no non trivial convergent sequences. The set $\{f_g\}$ defined above is then an example of a $w^*(E^*)$ compact set which contains no non trivial $w^*(E^*)$ sequential limits.

10. Reflexivity. The results in this section are consequences of the foregoing work and involve reflexivity of the Banach space E.

THEOREM 20. *If E is reflexive then E has the extended Helly property for all orders.*

This is an immediate consequence of Theorem 15 and the fact that reflexivity implies $w(E)$ compactness for S[16].

THEOREM 21. *A necessary and sufficient condition for E to have the extended Helly property is that E be reflexive.*

According to a result of Goldstine's [17], $w(E)$ completeness implies reflexivity. The present theorem thus follows from Theorems 16 and 20.

COROLLARY. *If E has the extended Helly property of order $\aleph_a$ and E^* contains $w^*(E^*)$ c-dense set of power $\aleph_a$ then E is reflexive.*

This is an immediate consequence of the corollary to Theorem 16 and Goldstine's theorem quoted above.

THEOREM 22. *If* (a) *S is $w(E)$ compact of order $\aleph_a$ and* (b) *E^* contains a $w^*(E^*)$ c-dense set of power $\leq \aleph_a$, then E is reflexive.*[8]

Let $\boldsymbol{U}_{\theta^*} = \{U^\mu(\theta^*) \mid \mu \in \mathfrak{A}\}$ be a $w^*(E^*)$ neighborhood basis at θ^* with $\mu \succ \mu' \Longleftrightarrow U^{\mu'} \supset U^\mu$. The basis $\boldsymbol{U}_\gamma$ at f_γ is obtained by a translation of $\boldsymbol{U}_{\theta^*}$ *preserving order,* of amount f_γ. Let $\{f^\iota\}$ be the $w^*(E^*)$ c-dense set of power $\aleph_a$. With each $U^\mu \in \boldsymbol{U}_\gamma$ associate an element $f^{\gamma,\mu} \in \{f^\iota\}$. By condition (b) such elements exist and may be chosen so as to be uniformly bounded. Then $\{f^{\gamma,\mu} \mid \mu \in \mathfrak{A}\} \rightarrow f_\gamma$ in the sense of $w^*(E^*)$ convergence.

Consider the family $\{A_a\}$ of $w(S)$ closed sets with the F. I. P. By Lemma 2, $A_a = \bigcap_\rho \bigcup_{i=1}^{2k} K_{i,\rho,a}$, $K_{i,\rho,a} = H_{i,\rho,a} \cap S$ where k depends on ρ and a. Hence, by Lemma 1, we can select a family $\{K_{i(\rho,a),\rho,a}\}$ with the F. I. P. In the interests of simplification we replace the ordering according to ρ and a by an ordering according to γ and write K_γ for $K_{i(\rho,a),\rho,a}$. The crux of the proof lies in establishing that $\bigcap_a A_a \neq \emptyset$.

---

[8] The topological nature of the proof is perhaps of greater interest than the theorem itself. The writer has recently observed that an alternative proof follows easily from some results of Alaoglo on characters of certain subspaces [2].

In view of Theorems 1 and 2, $K_\gamma = \bigcap_{\bar{\mu}} \bigcup_{j=1}^{N_{\gamma,\bar{\mu}}} B_{\gamma,j,\bar{\mu}}$. We apply Lemma 1 again and obtain the family $\{B_{\gamma,\mu}\}$ with the F. I. P. where

$$B_{\gamma,\bar{\mu}} = B_{\gamma,j(\gamma,\bar{\mu}),\bar{\mu}} = \{x \mid f^{\gamma,\mu(j(\gamma,\bar{\mu}),\bar{\mu})}(x) \geq f_\gamma(x_\gamma) + \epsilon_\gamma - \epsilon_{\bar{\mu}}\} \cap S.$$

The various values of $\mu(j(\gamma,\bar{\mu}),\bar{\mu})$, for which the associated $f^{\gamma,\mu(j(\gamma,\bar{\mu}),\bar{\mu})}$ is the same element of the well ordered collection $\{f^l\}$, are identified and their succession is determined by their position in $\{l\}$. They are now given a natural ordering by ν, $\nu = 1, 2 \cdots < \omega_a$, preserving this succession.

We well-order the pairs $(\gamma, \bar{\mu})$ and denote the typical ordered pair by w. Let $d_w = f_\gamma(x_\gamma) + \epsilon_\gamma - \epsilon_{\bar{\mu}}$. For each ν there is, in general, a multiplicity of values of w, indicated by writing $w[\nu]$, for which $\{x \mid f^\nu(x) \geq d_w\}$ is one of the sets $B_{\gamma,\bar{\mu}}$. Hence the sets $B_{\nu,w[\nu]} = \{x \mid f^\nu(x) \geq d_{w[\nu]}\} \cap S$ constitute (except for possible changes in order) the family $\{B_{\gamma,\bar{\mu}}\}$.

Let $d_{\nu_1} = \limsup \{d_{w[\nu_1]}\}$. Suppose provisionally that

$$d_{\nu_1} > d_{w[\nu_1]} \text{ for all values of } w[\nu_1]. \tag{22}$$

We select a monotone increasing sequence $\{d_{w_m[\nu_1]}\} \subset \{d_{w[\nu_1]}\}$ which converges to d_{ν_1}. Manifestly

$$B_{\nu_1,w'[\nu_1]} \supset B_{\nu,w[\nu_1]} \Longleftrightarrow d_{w[\nu_1]} \geq d_{w'[\nu_1]}.$$

Thus

$$B_{\nu_1,w_m[\nu_1]} = \bigcap\{w[\nu_1] \mid d_{w[\nu_1]} \leq d_{w_m[\nu_1]}\} B_{\nu_1,w[\nu_1]},$$

and more generally

$$C_{\nu_1} = \bigcap_{w[\nu_1]} B_{\nu_1,w[\nu_1]} = \bigcap_m B_{\nu_1,w_m[\nu_1]}.$$

Manifestly $C_{\nu_1} \neq \phi$ for $\{B_{\nu_1,w[\nu_1]}\}$ is a monotone decreasing sequence of $w(S)$ closed sets in the $w(E)$ compact, of order $\aleph_a \geq \aleph_0$, sphere S. If (22) is not valid, $d_{w[\nu_1]}$ assumes a maximum value, say for $w = w'$. Then $C_{\nu_1} = B_{\nu_1,w'[\nu_1]}$.

We now establish that $\{C_\nu\}$ has the F. I. P. Let ν be given the finite set of values $\nu_1, \cdots \nu_s$. Consider the families $\{B_{\nu_1,w[\nu_1]}\}, \cdots \{B_{\nu_s,w[\nu_s]}\}$. Suppose that (22) holds for $\nu_1, \cdots \nu_s$. (The trivial modifications when this is not true for some or all of $\nu_1, \cdots \nu_s$ are obvious from the remarks in the previous paragraph). For each i let $\{d_{w_m[\nu_i]}\}$ be the monotone ascending sequence defined, as before, which converges to d_{ν_1}. We write

$$B_m = B_{\nu_1,w_m[\nu_1]} \cap \cdots \cap B_{\nu_s,w_m[\nu_s]}.$$

Since $\{B_{\nu,w[\nu]}\}$ has the F. I. P., $B_m \neq \emptyset$. Moreover $\{B_m\}$ is plainly a monotone decreasing sequence of $w(S)$ closed sets. Hence, by the compactness condition (a),

$$\emptyset \neq \bigcap_m B_m = \bigcap_m \bigcap_{i=1}^{s} B_{\nu_i, w_m[\nu_i]} = \bigcap_{i=1}^{s} C_{\nu_i}.$$

Thus $\{C_\nu\} \subset S$, is a family of $w(S)$ closed sets with the F. I. P. and since $1 \leqq \nu < w_\alpha$ the $w(E)$ compactness of order $\aleph_\alpha$ of S implies

$$\emptyset \neq \bigcap_\nu C_\nu$$

or

$$\emptyset \neq \bigcap_\nu \bigcap_{w[\nu]} B_{\nu, w[\nu]} = \bigcap_{\gamma, \bar{\mu}} B_{\gamma, \bar{\mu}} \subset \bigcap_\gamma K_\gamma \subset \bigcap_\alpha A_\alpha.$$

We have established that any family $\{A_\alpha\}$ of $w(S)$ closed sets in S with the F. I. P. has a non-empty intersection. Therefore S is $w(E)$ compact and this implies that E is reflexive [16]. The proof is complete. (From the fact that $B_{\nu, w[\nu]}$ and C_ν are easily shown to be sections, it is evident that S need only be $w(E)$ *sectionally* compact of order $\aleph_\alpha$).

COROLLARY. *If* (a) *S is $w(E)$ countably compact and* (b) *E^* is $w^*(E^*)$ separable, then E is reflexive.*

This is a slight generalization of a theorem of Ghantmacher and Smulian [17] where the conditions are $w(E)$ sequential compactness of S and separability of E.

COROLLARY. *If* (a) *S^* is $w(E^*)$ countably compact and* (b) *E^* is $w^*(E^*)$ separable then E is reflexive.*

By a simple extension of a theorem of Ghantmacher-Smulian [17, Theorem 4] condition (a) is equivalent to condition (a) of the previous corollary.

Following Milman [11] we shall understand by a *face* corresponding to f, the point set $\{x \mid \| x \| = 1, f(x) = \| f \|\}$.

COROLLARY. *If* (a) *S is $w(E)$ countably compact and* (b) *the closed linear manifold M_f determined by a face of S always has a w^* separable conjugate space, then E is reflexive.*

Indeed the unit sphere of M_f is $S_f = M_f \cap S$. Now as a by-product of the analysis in Theorem 22, S_f is $w(M_f)$ compact. Since E is a Hausdorff space in the $w(E)$ topology it follows that S_f is absolutely closed in E and accordingly is $w(E)$ compact. Evidently the face corresponding to f is a strongly closed convex subset of S and so is also $w(E)$ compact and hence by Theorem 7 is tranfinitely closed. This fact together with (a) may be shown to imply that E is reflexive by the same reasoning as given by Milman for his theorem.

UNIVERSITY OF ILLINOIS,
ON LEAVE AT
THE INSTITUTE FOR ADVANCED STUDY.

BIBLIOGRAPHY

1. S. Banach, *Théorie des Operations Linéaires* (1932).

2. L. Alaoglu, "Weak topologies of normed linear spaces," *Annals of Mathematics*, vol. 41 (1940), pp. 252-267.

3. S. Kakutani, "Weak topology and regularity of Banach spaces," *Proceedings of the National Academy of Sciences of Tokyo*, vol. 15 (1939), p. 169.

3′. ———, "Weak topology, bicompact set and the principle of duality," *ibid.*, vol. 26 (1940), p. 62.

4. J. von Neumann, "Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren," *Mathematische Annalen*, vol. 102 (1929), pp. 370-427.

5. J. W. Tukey, "Convergence and uniformity in topology," *Annals of Mathematics Studies* (1940).

6. G. Birkhoff, "Moore-Smith convergence in general topology," *Annals of Mathematics*, vol. 38 (1937), pp. 39-56.

7. H. Wallman, "Lattices and topological spaces," *Annals of Mathematics*, vol. 39 (1939), pp. 112-126.

8. J. W. Alexander, "Ordered sets, complexes and the problem of compactification," *Proceedings of the National Academy of Sciences*, vol. 25 (1939), pp. 296-298.

9. P. Alexandroff and H. Hopf, *Topologie*, 1, Berlin (1935).

10. G. Sirvint, "Schwache Kompaktheit in den Banachschen Raumen," *C. R. URSS*, vol. 28 (1940), pp. 199-202.

11. D. Milman, "On some criteria for the regularity of spaces of the type (B)," *C. R. URSS*, vol. 20 (1938), pp. 243-246.

12. E. Helly, "Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten," *Monatshefte für Math. und Physik*, vol. 31 (1921), pp. 60-91.

13. S. Mazur, "Über konvexe Mengen in linearen normierten Räumen," *Studia Mathematica*, vol. 4 (1933), pp. 70-84.

14. M. Krein and V. Smulian, "Regularly convex sets," *Annals of Mathematics*, vol 41 (1940), pp. 556-583.

15. A. Tychonoff, "Über einen Funktionenraum," *Mathematische Annalen*, vol. 111 (1935), pp. 762-766.

16. N. Bourbaki, "Sur les espaces de Banach," *Comptes Rendus*, vol. 206 (1938), pp. 1701-1704.

17. H. H. Goldstine, "Weakly complete Banach spaces," *Duke Mathematical Journal*, vol. 4 (1938), pp. 125-131.

18. V. Ghantmacher and V. Smulian, "Sur les espaces linéaires dont la sphère unitaire est faiblement compacte," *C. R. URSS*, vol. 17 (1937), pp. 91-94.

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

CALCUTTA CENTRAL LIBRARY UNIVERSITY

UNIFORM LOCAL CONNECTEDNESS AND CONTRACTIBILITY.*

By SAMUEL EILENBERG and R. L. WILDER.

Although we shall be concerned chiefly with uniform local connectedness in the sense of homotopy, we shall make contact frequently with the analogous notions as defined in terms of homology and contractibility. The definitions of these and other concepts are given in Section **1**.

In Sections **2-5** we investigate the properties of uniformly locally connected subsets of a metric separable space with particular reference to the relation between the set and its closure. The substitution of local connectedness of a set *in relation to the space* is of course a device to avoid imposing compactness; in the compact case, as, for instance, when the imbedding space is the euclidean n-sphere S^n, the notion is equivalent to that of uniform local connectedness.

In Sections **5** and **6** applications are made to the subsets of euclidean space. We find, for example, that if a domain complementary to the homeomorph of an S^{n-1} in S^n is uniformly locally connected in the homotopy sense[1] in all dimensions 0 to n, then the domain is a singular n-cell.

1. Let X be a metric separable space, A a subset of X, and $f(Z) \subset X$ a continuous mapping of a space Z. Then we call f *nullhomotopic over* A if there is a mapping[2]

$$g(Z \times E) \subset A \qquad E = [0, 1] \tag{1}$$

such that

$$\begin{cases} g(z, 0) = f(z) \\ g(z, 1) = \text{constant} \end{cases} \quad \text{for every } z \, \varepsilon \, Z. \tag{2}$$

We also define [7] a number $b(f, A)$ as follows: $b(f, A) = \infty$ if f is not nullhomotopic over A; $b(f, A) = \text{g.l.b.}\, \delta[g(Z \times E)]$ if f is nullhomotopic over A and g varies over all mappings (1) satisfying (2). The number $b(f, A)$ may be called the *homotopy bound* of f over A.

Let Q^{n+1} be the $(n+1)$-cube and S^n its boundary. We easily verify the following lemma:

---

* Presented to the American Mathematical Society, April 26, 1940, under the title "Uniform local contractibility"; Received August 11, 1941.

[1] It is known that such a domain is always uniformly locally connected in the homology sense in all dimensions [8]. (We use numerals in brackets to refer to the Bibliography.)

[2] All mappings considered in this paper will be assumed continuous without explicit statement of the fact hereafter.

(3) *Given a mapping* $f(S^n) \subset X$ *and a subset* $A \subset X$, *we have* $b(f, A) < \epsilon$ *if and only if there is an extension*[3] $f^*(Q^{n+1}) \subset X$ of f such that $f^*(Q^{n+1}) \subset A$ *and* $\delta[f^*(Q^{n+1})] < \epsilon$.

Let $x \varepsilon X$ and let A be again a subset of X. We say that A is *locally p-connected in the homotopy sense at* x ($=$ *p-LC* at x) if for every $\epsilon > 0$ there is an $\eta > 0$ such that every mapping $f(S^p) \subset A \cdot S(x, \eta)$[4] is nullhomotopic in $A \cdot S(x, \epsilon)$. Clearly A is *p-LC* at x if $x \varepsilon X - \bar{A}$. We say that A is *locally p-connected in the homotopy sense* (A is *p-LC*) if A is *p-LC* at every $x \varepsilon A$. If the same holds at every $x \varepsilon X$, then we say that A is *locally p-connected in relation to* X (A is *p-LC* rel. X). If A is *p-LC* for $p = 0, 1, \cdots, n$, then we say that A is LC^n. Similarly we define LC^n rel. X.

Using the number $b(f, A)$ introduced above, all these definitions can apparently be stated without the "ϵ, η" procedure. In fact, it is easy to see that A is *p-LC* at x whenever, given a sequence $f_k(S^p) \subset A$ of mappings such that $\operatorname{Lim}_k f_k(S^p) = x$, we have $\lim_k b(f_k, A) = 0$.

If A is *p-LC* and if, for any given $\epsilon > 0$, the choice of η does not depend on the point x, then we say that A is *uniformly locally p-connected in the sense of homotopy* (A is *p-ULC*). If the same holds for $p = 0, 1, \cdots, n$, we say that A is ULC^n. Using the number $b(f, A)$ we find that A is *p-ULC* if and only if, given $\epsilon > 0$, there is an $\eta > 0$ such that if $f(P^p) \subset A$ and $\delta[f(S^p)] < \eta$ then $b(f, A) < \epsilon$.

The following statement is easily verified:

(4) *If* A *is* ULC^n *then* A *is* LC^n *rel.* X. *If* $\bar{A}$ *is compact then the two concepts are equivalent.*

Concepts similar to the above may be defined for local connectedness in the sense of homology using Vietoris cycles and chains[5] on compact carriers. In the latter case we shall use lower case letters in the corresponding symbols; for instance, locally p-connected in the homology sense will be designated by the symbol p-lc.

Finally, we shall say that A is *locally contractible at* $x \epsilon X$ (A is $\mathcal{LC}$ at x) if, given $\epsilon > 0$, there is an $\eta > 0$ such that $A \cdot S(x, \eta)$ is contractible in $A \cdot S(x, \epsilon)$ to a point. Definitions of "A is $\mathcal{LC}$," "A is $\mathcal{LC}$ rel. X" and "$\mathcal{ULC}$" are obtained by replacing "*p-LC*" in the corresponding definitions above by "$\mathcal{LC}$." Clearly A is $\mathcal{ULC}$ if and only if for every $\epsilon > 0$ there is an $\eta > 0$ such that every subset of A of diameter $< \eta$ is contractible to a point over a subset of A of diameter $< \epsilon$. Corresponding to (4) we have:

---

[3] Given a subset B of a space Z and a mapping $f(B) \subset X$, then a mapping $f^*(Z) \subset X$ will be called an extension of f if $f^*(z) = f(z)$ for $z \varepsilon B$.

[4] By $S(x, \eta)$ we denote the set of all points y such that $\rho(x, y) < \eta$.

[5] It should be recalled here that the "lc" concepts depend on the abelian group chosen as coefficient domain. Where it is essential to do so, we designate this group.

(5) *If A is $\mathcal{ULC}$ then it also is $\mathcal{LC}$ rel. X. If $\bar{A}$ is compact then the two concepts coincide.*

2. Throughout this section X will denote a metric separable space and A a subset of X.

THEOREM 1. *A is LC^n rel. X if, and only if, given any closed subset B of a metric separable space Z such that $\dim(Z - B) \leqq n + 1$, and given a mapping $f(B) \subset \bar{A}$, there is an open set $U \supset B$ and an extension*[8] *$f'(U) \subset \bar{A}$ of f such that $f'(U - B) \subset A$.*

Proof of necessity. We may admit that $\bar{A} = X$.

According to a theorem of Kuratowski [6, Th. 2] X can be imbedded in a metric separable space Y such that

(1) X is a closed subset of Y;

(2) $Y - X = P^{n+1}$ is an infinite polyhedron of dimension $\leqq n + 1$;

(3) the mapping $f(B) \subset X$ has an extension $y(Z) \subset Y$ such that $y(Z - B) \subset P^{n+1}$.

In view of (3) it is therefore sufficient to prove that

(4) there is an open set V such that $X \subset V \subset Y$ and a mapping $r(V) \subset X$ such that $r(x) = x$ for $x \varepsilon X$ and $r(V - X) \subset A$. We proceed to construct the set V and mapping r.

Let P^k denote the k-dimensional skeleton of P^{n+1}; i. e., the subpolyhedron of P^{n+1} consisting only of the simplices of dimension $\leqq k$. To prove (4) it is sufficient to prove that for $k = 0, 1, \cdots, n + 1$,

(5) there is an open set V_k such that $X \subset V_k \subset Y$ and a mapping $r_k(X + V_k \cdot P^k) \subset X$ such that $r_k(x) = x$ for $x \varepsilon X$, $r_k(V_k \cdot P^k) \subset A$. Hence we may proceed by induction. For $k = 0$, P^0 is the set of all vertices of P^{n+1}. Since A is dense in $X = \bar{A}$, we can find for every vertex $x_i \varepsilon P^0$ a point $y_i \varepsilon A$ such that $\rho(x_i, y_i) < 2\rho(x_i, X)$. Defining $V_0 = Y$ and $r_0(x) = x$ for $x \varepsilon X$, $r_0(x_i) = y_i$ for $x_i \varepsilon P^0$, we verify (5) for $k = 0$.

Suppose that V_k and r_k of (5) are given for some $k < n + 1$. Let $\Delta_1^{k+1}, \Delta_2^{k+1}, \cdots$, be the $(k+1)$-dimensional simplexes contained in $V_k \cdot P^{n+1}$. Denote by S_i^k the boundary sphere of Δ_i^{k+1}. The mapping $r_k(S_i^k) \subset A$ is therefore defined. Denote this partial mapping by f_i: $f_i(S_i^k) \subset A$. If $b(f_i, A) < \infty$ then we can find a mapping $g_i(\Delta_i^{k+1}) \subset A$ such that $g_i(x) = f_i(x) = r_k(x)$ for $x \varepsilon S_i^k$ and such that $\delta[g_i(\Delta_i^{k+1})] < 2b(f_i, A)$. Now if the subsequence $\{\Delta_{i_s}^{k+1}\}$ converges to a point $a \in X$ then $\delta[r_k(S_{i_s}^{k})] \to 0$ and $\delta[f_{i_s}(S_{i_s})] \to 0$; and since A is k-LC at a we must have $b(f_{i_s}, A) \to 0$ and therefore $g_{i_s}(\Delta_{i_s}^{k+1})$ converges to a.

Consequently there is an open set V_{k+1} such that $X \subset V_{k+1} \subset Y$ and $b(f_i, A)$ is finite whenever $\Delta_i^{k+1} \cdot V_{k+1} \neq 0$. Taking

$$r_{k+1}(x) = r_k(x) \text{ for } x \in X + V_{k+1} \cdot P^k; \quad r_{k+1}(x) = g_i(x) \text{ for } x \in V_{k+1} \cdot \Delta_i^{k+1},$$

we verify (5) for $k + 1$.

Proof of sufficiency. Suppose that at some $a \in X$ the set A is not p-LC for some $p \leqq n$. There is, therefore, a sequence $f_i(S^p) \subset A$ of mappings such that $\mathrm{Lim}_i\, f_i(S^p) = a$ and $b(f_i, A) > \epsilon$ for some $\epsilon > 0$.

Consider a space Z consisting of a sequence of disjoint $(p+1)$-cubes $\{Q_i^{p+1}\}$ and a point z such that $z = \mathrm{Lim}_i\, Q_i^{p+1}$. In Z consider the subset $B = z + \Sigma_i S_i^p$, where S_i^p is the boundary sphere of Q_i^{p+1}. The sequence $\{f_i\}$ clearly defines a mapping $f(B) \subset \bar{A}$ such that $f(z) = a$ and $f(B - z) \subset A$. Since $\dim(Z - B) \leqq n + 1$ there is an open set U such that $B \subset U \subset Z$ and an extension $f^*(U) \subset \bar{A}$ of f such that $F^*(U - B) \subset A$. For sufficiently large i we then have $f^*(Q_i^{p+1}) \subset A$ and $\delta[f^*(Q_i^{p+1})] < \epsilon$. This contradicts $b(f_i, A) > \epsilon$.

Remark. If $\bar{A}$ is compact then Section 1, (4) and Theorem 1 give a necessary and sufficient condition for A to be ULC^n.

If $A = X$ then Theorem 1 furnishes the following theorem due to Kuratowski [7, Th. 1]:

THEOREM 1a. *X is LC^n if, and only if, given any closed subset B of a metric separable space Z such that $\dim(Z - B) \leqq n + 1$, and given a mapping $f(B) \subset X$, there is an open set $U \supset B$ and an extension $f'(U) \supset X$ of f.*

Both Theorems 1 and 1a can be restated for Z metric compact, for the space Z used in the sufficiency proof was compact metric.

THEOREM 2. *Let B be a closed subset of a compact space Z such that $\dim(Z - B) \leqq n + 1$. If A is LC^n rel. X then given a mapping $f(Z) \subset \bar{A}$ there is a mapping $f^*(Z) \subset \bar{A}$ such that $f^*(Z - B) \subset A$ and $f^*(z) = f(z)$ for $z \in B$.*

Moreover, given any $\epsilon > 0$, we can choose f^* so that $\rho[f(z), f^*(z)] < \epsilon$ for every $z \in Z$.

Proof. Consider in the product space $Z \times E$ $(E = [0, 1])$ the sets

$$Z^* = Z \times 0 + \sum_{n=1}^{\infty} Z \times (1/n); \qquad B^* = Z \times 0 + \sum_{n=1}^{\infty} B \times (1/n).$$

Clearly $B^* \subset Z^*$ and $\dim (Z^* - B^*) \leqq n + 1$ since $\dim (Z - B) \leqq n + 1$. Define the mapping $g(B^*) \subset \bar{A}$ as follows:

$$g(z, 0) = f(z) \text{ for } z \,\varepsilon\, Z; \qquad g(z, 1/n) = f(z) \text{ for } z \,\varepsilon\, B.$$

Since A is LC^n rel. X we may find, by Theorem 1, an open set U such that $B^* \subset U \subset Z^*$ and a mapping $g^*(U) \subset \bar{A}$ such that $g^*(x) = g(x)$ for $x \,\varepsilon\, B^*$ and $g^*(U - B^*) \subset A$. For sufficiently large n we then have $Z \times (1/n) \subset U$; and taking $f_n(z) = g^*(z, 1/n)$ we have $f_n(Z) \subset \bar{A}$, $f_n(Z - B) \subset A$ and $f_n(z) = f(z)$ for $z \,\varepsilon\, B$; and, moreover, $\rho[f_n(z), f(z)] < \epsilon$.

Remark. If $\bar{A}$ is LC^n then the conclusion of Theorem 2 is a necessary and sufficient condition for A to be LC^n rel. X.

The following theorem is an immediate consequence of Theorem 2:

THEOREM 3. *Let Z be a compact space such that $\dim Z \leqq n$. If A is LC^n rel. X then every two mappings $f_1(Z) \subset A$ and $f_2(Z) \subset A$ which are homotopic in $\bar{A}$ are also homotopic in A.*

THEOREM 4. *If A is LC^n rel. X then so is every set A^* such that $A \subset A^* \subset \bar{A}$. In particular $\bar{A}$ is LC^n.*

To establish Theorem 4 it is sufficient to note that if the condition of Theorem 1 holds for A then it holds *a fortiori* for A^*.

Remark. Let us also remark that it easily follows from Theorems 2 and 3 that A and A^* have the same homotopy and homology groups for the dimensions $0, 1, \cdots, n$.

3. In this section we again assume that A is a subset of a metric separable space X, but with the additional assumption that $\dim \bar{A} \leqq n$.

THEOREM 5. *The following properties are equivalent:*

(a) *A is $\mathfrak{LC}$ rel. X.*
(b) *A is LC^n rel. X.*
(c) *A is LC^n and there is a mapping $h(\bar{A} \times E) \subset \bar{A}$ such that $h(x, 0) = x$, $h(x, t) \,\varepsilon\, A$ for $t > 0$.*
(The deformation h will be frequently used in the sequel.)

Proof. That (a) implies (b) is obvious.

(b) implies (c). Clearly if A is LC^n rel. X then A must also be LC^n (rel. A). We now proceed to construct the mapping h.

Define $Z = \bar{A} \times E$, $B = \bar{A} \times (0)$, and $f(x, 0) = x$ for $x \,\varepsilon\, \bar{A}$. We have then $B \subset Z$, $\dim (Z - B) \leqq n + 1$ and $f(B) \subset \bar{A}$. Since A is LC^n rel. X

there are, by Theorem 1, an open set U such that $B \subset U \subset Z$ and an extension $f^*(U) \subset \bar{A}$ of f such that $f^*(U - B) \subset A$.

Since X is metric, Z is metric, and we may define a function $r(x) = \frac{1}{2}\rho[(x,0), Z - U]$ for $x \varepsilon \bar{A}$. Clearly $r(x)$ is continuous and > 0. The mapping h will be defined as follows:

$$h(x,t) = f^*[x, t \cdot r(x)].$$

Since $0 \leqq t \cdot r(x) \leqq r(x)$, $h(x,t)$ is well defined and satisfies all the requirements of Theorem 5.

(c) implies (a). Since $\dim A \leqq n$ and A is LC^n, it follows that A is $\mathcal{LC}$ (rel. A) [7, § 8. This can be easily deduced from Theorem 1a of the present paper]. To prove that A is $\mathcal{LC}$ rel. X, consider a point $x \varepsilon \bar{A}$ and a sequence $\{A_i\}$ of subsets of A such that $\operatorname{Lim}_i A_i = x$.

Given any $\epsilon > 0$ there is an open set U containing x and an $\eta > 0$ such that if $y \varepsilon \bar{A} \cdot U$ and $0 \leqq t \leqq \eta$ then $\rho[x, h(y,t)] < \epsilon$. For sufficiently large i we must have $\delta(A_i) < \epsilon$ and $A_i \subset U$. It follows that

$$\delta[\sum_{0 \leqq t \leqq \eta} h(A_i, t)] < 3\epsilon$$

and therefore A_i is deformed into $h(A_i, \eta)$ over a subset of A of diameter $< 3\epsilon$. Because of continuity we have $\operatorname{Lim}_i h(A_i, \eta) = h(x, \eta)$, and since $h(x,\eta) \varepsilon A$ and A is $\mathcal{LC}$ it follows that for sufficiently large i the set $h(A_i, \eta)$ can be contracted to a point over a subset of A of diameter $< \epsilon$. Combining these two deformations we see that for sufficiently large i the set A_i can be contracted to a point over a subset of A of diameter $< 4\epsilon$. This proves that A is $\mathcal{LC}$ rel. X.

THEOREM 6. *If A is $\mathcal{LC}$ rel. X and $A \subset A^* \subset \bar{A}$ then* (a) *A^* is $\mathcal{LC}$ rel. X and* (b) *A and A^* have the same homotopy type* [4].

The first part of Theorem 6 is an immediate consequence of Theorems 4 and 5. In order to establish the second part define the mappings $g_1(A) \subset A^*$, $g_2(A^*) \subset A$ as follows:

$$g_1(x) = x \text{ for } x \varepsilon A; \qquad g_2(x) = h(x,1) \text{ for } x \varepsilon A^*,$$

where h is the mapping of Theorem 5. It follows from the properties of h that g_1g_2 is homotopic to the identity mapping of A^* into itself and that g_2g_1 is homotopic to the identity mapping of A into itself.

THEOREM 7. *If A is LC^n rel. X and A is contractible in $\bar{A}$ then there is a mapping $g(\bar{A} \times E) \subset \bar{A}$ such that $g(x,0) = x, g(x,t) \varepsilon A$ for $t > 0$, and $g(x,1) = constant$.*

Proof. Let $f(A \times E) \subset \bar{A}$ be a mapping such that $f(x,0) = x$ and $f(x,1) =$ constant. Then, h being the mapping of Theorem 6, let

$$g(x,t) = h[h(x,t),t] \text{ for } x \,\varepsilon\, \bar{A}, 0 \leqq t \leqq \tfrac{1}{2},$$
$$g(x,t) = h\{f[h(x,t),2t-1],t\} \text{ for } x \,\varepsilon\, \bar{A}, \tfrac{1}{2} \leqq t \leqq 1.$$

4. In this section we shall assume that A is an *open* subset of a metric separable space X. We shall denote by M the boundary of A; $M = \bar{A} - A$.

LEMMA. *Suppose that M is LC^1 and that every mapping $f(S^1) \subset M$ is nullhomotopic in $\bar{A}$. If $C \subset \bar{A}$ is a set such that every mapping $f(S^1) \subset C$ is nullhomotopic in X then every mapping $f(S^1) \subset C$ is also nullhomotopic in $\bar{A}$.*

Proof. Let Q^2 be a closed disc bounded by S^1. Consider a mapping $f(S^1) \subset C$. Since f is nullhomotopic in X there is an extension $f_1(Q^2) \subset X$ of f.

Consider the subsets $B_1 = f_1^{-1}(M)$ and $B = f_1^{-1}(\bar{A})$ of Q^2. Clearly $B_1 \subset B$, $S_1 \subset B$, and B_1, B are closed subsets of Q^2. Also consider the mapping $f_1(B_1) \subset M$. Since M is LC^1 and $\dim (Q^2 - B_1) \leqq 2$, there are, by Theorem 1a, an open set U such that $B_1 \subset U \subset Q^2$ and a mapping $f_2(U) \subset M$ such that $f_1(q) = f_2(q)$ for $q \,\varepsilon\, B_1$. Define

$$V = B + U; \quad f_3(q) = f_1(q) \text{ for } q \,\varepsilon\, B; \quad f_3(q) = f_2(q) \text{ for } q \,\varepsilon\, U - B.$$

The set V is obviously open and $S^1 \subset B \subset V$. The mapping $f_3(V) \subset \bar{A}$ is continuous and has the following properties:

$$f_3(q) = f(q) \text{ for } q \,\varepsilon\, S^1; \qquad f_3(q) \,\varepsilon\, M \text{ for } q \,\varepsilon\, V - B.$$

Since B is a closed subset of Q^2 containing its boundary and V is an open set containing B there is in Q^2 a finite set of disjoint discs $Q_1^2, Q_2^2, \cdots, Q_n^2$ bounded by simple closed curves $S_1^1, S_2^1, \cdots, S_n^1$ such that

$$Q_1^2 + Q_2^2 + \cdots + Q_n^2 + V = Q^2; \quad S_1^1 + S_2^1 + \cdots + S_n^1 \subset V - B.$$

Since $S_i^1 \subset V - B$ we have $f_3(S_i^1) \subset M$, and since this mapping must be nullhomotopic in $\bar{A}$ there is a mapping $g_i(Q_i^2) \subset \bar{A}$ such that $f_3(q) = g_i(q)$ for $q \,\varepsilon\, S_i^1$. If we define

$$f_4(q) = f_3(q) \text{ for } q \,\varepsilon\, Q^2 - (Q_1^2 + Q_2^2 + \cdots + Q_n^2);$$
$$f_4(q) = g_i(q) \text{ for } q \,\varepsilon\, Q_i^2,$$

we obtain a mapping $f_4(Q^2) \subset \bar{A}$ such that $f_4(q) = f(q)$ for $q \,\varepsilon\, S^1$. This proves that f is nullhomotopic in $\bar{A}$.

THEOREM 8. *If* (a) *M is LC^1,* (b) *A is LC^1 rel. X,* (c) *every mapping $f(S^1) \subset M$ is nullhomotopic in $\bar{A}$ and* (d) *every mapping $f(S^1) \subset A$ is null-*

homotopic in X, *then* $\pi_1(A) = 0$,[6] *and more generally* $\pi_1(A^*) = 0$ *whenever* $A \subset A^* \subset \bar{A}$.

Indeed, let $f(S^1) \subset A^*$. Applying the Lemma for $C = A$, we find that f is nullhomotopic in $\bar{A}$. Since A^* is LC^1 rel. X it follows from Theorem 4 that f is nullhomotopic in A^*.

As a corollary of Theorem 8 we obtain

THEOREM 9. *If* (a) M *is* LC^1, (b) A *is* LC^1 *rel.* X, (c) $\pi_1(M) = 0$, (d) $\pi_1(X) = 0$ *and* (e) $A \subset A^* \subset \bar{A}$, *then* $\pi_1(A^*) = 0$.

5. Although the previous theorems were proved for quite general spaces, their intuitive geometric content becomes clearer if they are applied to euclidean spaces. In this section we shall assume that A is an open subset of the euclidean n-sphere S^n. As before we shall denote by M the boundary $\bar{A} - A$ of A.

Since $\bar{A}$ is compact every local connectedness property of A rel. S^n is a uniform local connectedness property of A (§ 1). From Theorem 5 we therefore get

THEOREM 5a. *The following properties are equivalent:*

(a) A *is* $\mathcal{ULC}$,

(b) A *is* ULC^n,

(c) *there is a mapping* $h(\bar{A} \times E) \subset \bar{A}$ *such that* $h(x, 0) = x$, *and* $h(x, t) \varepsilon A$ *for* $t > 0$.

Following R. L. Wilder [9, "Δ-deformable"] we say that M is *deformation-free into* A if there is a mapping $f(M \times E) \subset \bar{A}$ such that $f(x, 0) = x$, and $f(x, t) \varepsilon A$ for $t > 0$. It follows from Theorem 5a that if A is ULC^n then M is deformation-free into A. On the other hand it was proved by R. L. Wilder [9, 10] that if such an M is deformation-free into A then A is ulc^n and consequently M is the sum of a finite number of disjoint generalized $(n-1)$-manifolds.

PROBLEM 1. *If* M *is a continuum separating* S^n, *is it true that* M *is deformation-free into* A *if and only if* A *is* ULC^n ($= \mathcal{ULC}$ *since* $A \subset S^n$)?

Since $\pi_1(S^n) = 0$,[6] Theorem 8 gives the following

THEOREM 8a. *If* M *is* LC^1, A *is* ULC^1 *and every mapping* $f(S^1) \subset M$

---

[6] $\pi_1(A)$ denotes the fundamental group of A, and the relation $\pi_1(A) = 0$ means that every mapping $f(S^1) \subset A$ is nullhomotopic.

is nullhomotopic in $\bar{A}$, *then* $\pi_1(A^*) = 0$ *whenever* $A \subset A^* \subset \bar{A}$. *In particular* $\pi_1(A) = 0 = \pi_1(\bar{A})$.

THEOREM 9a. *If* M *is* LC^1, A *is* ULC^1 *and* $\pi_1(M) = 0$, *then* $\pi_1(A^*) = 0$ *whenever* $A \subset A^* \subset \bar{A}$. *In particular* $\pi_1(A) = 0 = \pi_1(\bar{A})$.

We do not know if the condition that M is LC^1 is necessary. In general, as may be shown by examples, A may be ULC^1 without M being LC^1. However, there is the following

PROBLEM 2. *Find conditions which* A *must satisfy in order that* M *be* LC^k. *And if* A *is* ULC^n, *is* M *an* LC^n? [7]

We remark here that R. L. Wilder [11] has found necessary and sufficient conditions on A for M to be lc^k over a finite coefficient group. For $k = 0$ all classes lc^0 and LC^0 coincide, so that for $k = 0$ a solution of Problem 2 follows from Wilder's theorem. For $k > 0$, we know from a theorem of Hurewicz [5] that M is LC^k if and only if it is lc^k over the ring of integers and LC^1.

6. In this section we assume that $M \subset S^n$ is a homeomorph of S^{n-1} and that A is one of the two components of $S^n - M$. An example was constructed by Alexander [1] to show that $\pi_1(A)$ [6] does not necessarily vanish. We shall see from the following theorem that this cannot be the case if A is ULC^1.

THEOREM 10. *If* A *is* ULC^1 *then both* A *and* $\bar{A}$ *are contractible and* $\bar{A}$ *is an absolute retract.*

Since M is an absolute neighborhood retract [2] it is clear that $\bar{A} = A + M$ is one too. Since A is ULC^1 it follows from Theorem 8a that $\pi_1(A) = 0 = \pi_1(\bar{A})$. Since also all the homology groups of A and $\bar{A}$ vanish, it follows from a theorem of Hurewicz [3] that both A and $\bar{A}$ are contractible.

THEOREM 11. A *is* ULC^n *if and only if there is a mapping* $h(\bar{A} \times E) \subset \bar{A}$ *such that* $h(x, 0) = x$, $h(x, 1) =$ *constant, and* $h(x, t) \in A$ *for* $t > 0$.

In fact if A is ULC^n then by virtue of Theorem 10, A is contractible and the existence of h is guaranteed by Theorem 7. If a mapping h exists satisfying the conditions of Theorem 11, then A is ULC^n because of Theorem 5a.

Let Q^n be the n-dimensional cube and S^{n-1} its boundary. Since S^{n-1} and M are homeomorphic the following theorem is an immediate consequence of Theorem 11 and elementary properties of linked cycles:

---

[7] Since $A \subset S^n$, the second part of Problem 2 can be reformulated as follows: If A is $\mathcal{ULC}$, is M an $\mathcal{LC}$ (= absolute neighborhood retract)?

THEOREM 12. *If* A *is* ULC^n *then there is a mapping* $f(Q^n) = \bar{A}$ *such that* $f(Q^n - S^{n-1}) = A$ *and which maps* S^{n-1} *homeomorphically onto* M.

PROBLEM 3. *Is the converse of Theorem* 12 *true?* It is easy to see that a positive answer would follow from a positive solution of Problem 1.

Since M is a manifold it follows from the results of R. L. Wilder [8] that A is ulc^n. It follows from the example of Alexander [1] and from Theorem 10 that A may not be ULC^1. We do not know whether A can be ULC^1 without being ULC^n. This question is part of the following more general

PROBLEM 4. *Let* A *be an open subset of a compact metric space. Is it true that* A *is* ULC^n *if and only if it is* ULC^1 *and* ulc^n *over the ring of integers?*[8]

UNIVERSITY OF MICHIGAN.

BIBLIOGRAPHY

1. J. W. Alexander, "An example of a simply connected surface bounding a region which is not simply connected," *Proceedings of the National Academy of Sciences*, vol. 10 (1924), pp. 8-10.

2. K. Borsuk, "Über eine Klasse von lokal zusammenhängenden Räumen," *Fundamenta Mathematicae*, vol. 19 (1932), pp. 220-242.

3. W. Hurewicz, "Beiträge zur Topologie der Deformationen," II, *Koninklijke Akademie van Wetenschappen te Amsterdam, Proceedings*, vol. 38 (1935), pp. 521-528.

4. ———, "Beiträge zur Topologie der Deformationen," III, *ibid.*, vol. 39 (1936), pp. 117-125.

5. ———, "Homotopie, Homologie und lokaler Zusammenhang," *Fundamenta Mathematicae*, vol. 25 (1935), pp. 467-485.

6. C. Kuratowski, "Sur le prolongement des fonctions continues et les transformations en polytopes," *Fundamenta Mathematicae*, vol. 24 (1935), pp. 259-268.

7. ———, "Sur les espaces localement connexes et péaniens en dimension n," *ibid.*, pp. 269-287.

8. R. L. Wilder, "Generalized closed manifolds in n-space," *Annals of Mathematics*, vol. 35 (1934), pp. 876-903.

9. ———, "Concerning a problem of K. Borsuk," *Fundamenta Mathematicae*, vol. 21 (1933), pp. 156-167.

10. ———, "On free subsets of E_n," *ibid.*, vol. 25 (1935), pp. 200-208.

11. ———, "Characterizations of those euclidean domains whose boundaries are lc^k," unpublished, abstracted in *Bulletin of the American Mathematical Society*, vol. 47 (1941), p. 58.

---

[8] If A is the whole space the answer is positive in view of Hurewicz's results [5].

ON THE UNIFIED THEORY OF EIGENVALUES OF PLATES AND MEMBRANES.*

By NATHAN ARONSZAJN and ALEXANDER WEINSTEIN.

---

1. Introduction. In a previous paper[1] one of the authors has shown that the problem of the vibrations of a membrane and the problem of the vibrations of a clamped (or partially clamped and partially supported) plate are two limiting cases of a sequence of intermediate problems. A similar relation holds between the membrane problem and the buckling problem for plates. The eigenvalues of these *intermediate problems* (which were called in *W* the *modified problems*) give a non-decreasing sequence of lower bounds for the eigenvalues of plates.

The consideration of these intermediate problems was of interest because, while the differential problems of the theory of plates are of the fourth order, the differential equations of the intermediate problems are of the second order; moreover their solutions can be expressed in terms of the solutions of the membrane problem.

In a subsequent note in common[2] we investigated the conditions for the convergence of the sequence of lower bounds to the exact eigenvalues of the plate. In the present paper we develop the proof of the convergence theorem already sketched in our note. Moreover we give proofs of the existence theorems for our problems and we discuss the question of the equivalence of the different formulations which can be given to the eigenvalue problems of the clamped or partially clamped plate.

For definiteness, we deal chiefly in the following with the problem of the vibrations of a clamped plate and we consider this problem from the variational point of view. The corresponding differential problem will be discussed in **6.** The case of a plate partially clamped and partially supported, which requires only slight modifications, will be treated in **8.** Finally, in **9,** the theory is adapted to the case of the buckling of a plate.

---

* Received May 25, 1941; an abstract of this paper, under the title, "Existence, convergence and equivalence in the unified theory of eigenvalues of plates and membranes" appeared in the *Proceedings of the National Academy of Sciences*, vol. 27 (1941), pp. 188-191.

[1] A. Weinstein, "Etude des spectres des équations aux dérivées partielles de la théorie des plaques élastiques," ***Mémorial des Sciences Mathématiques***, 88, Thèse, Paris (1937). This paper will be referred to as W. Cf. also A. Weinstein, "Sur la théorie unitaire des valeurs propres des membranes et des plaques encastrées," ***Comptes rendus de l'Académie des Sciences***, Paris, 210 (1940), p. 161.

[2] N. Aronszajn et A. Weinstein, "Sur la convergence d'un procédé variationnel d'approximation dans la théorie des plaques encastrées," ***Comptes rendus de l'Académie des Sciences***, 204 (1937), p. 96.

Let S be the domain of our plate in the (x, y)-plane. The boundary C of S is supposed to satisfy certain conditions of regularity which will later be formulated explicitly.

The following notations will be used throughout this paper:

$$H(w) = \int\int_S w^2 dxdy, \qquad I(w) = \int\int_S (\Delta w)^2 dxdy,$$

$$(v, w) = \int\int_S vw \, dxdy.$$

We have obviously

$$I(w) = H(\Delta w) = (\Delta w, \Delta w).$$

Using these notations we can formulate as follows the variational problem $I^{(n)}$ which gives the eigenvalue $\lambda^{(n)}$ $(n = 1, 2, \cdots)$ in the problem of the vibrating clamped plate.

PROBLEM $I^{(n)}$. *To find a function minimizing the expression* $I(w)/H(w)$ *among all functions* w *continuous in* $S + C$, *possessing continuous first and second derivatives in* S, *the Laplacian* Δw *being of integrable square in* S. *Furthermore* w *must satisfy the following boundary conditions*

$$(1.1) \qquad w = 0 \quad \text{on } C$$

$$(1.2) \qquad \lim w_x = \lim w_y = 0,$$

for (x, y) *tending to* C. *Moreover, for* $n > 1$, *we must satisfy the following* $n - 1$ *orthogonality conditions*

$$(1.3) \qquad (w, w^{(j)}) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1,$$

where the notation $w^{(i)}$ *is used in a general way for a solution of* $I^{(i)}$ $(i = 1, 2, \cdots)$. *The minimum in* $I^{(n)}$ *is called the eigenvalue* $\lambda^{(n)}$ *of the clamped plate.*

In line with the general theory of eigenvalues it will be useful to consider, for $n > 1$, a Problem $\hat{I}^{(n)}$ generalizing $I^{(n)}$. The formulation of $\hat{I}^{(n)}$ differs from that of $I^{(n)}$ only with respect to the conditions (1.3) for which we have to substitute the $n - 1$ orthogonality conditions

$$(1.4) \qquad (w, \phi_j) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1,$$

where $\phi_1, \phi_2, \cdots, \phi_{n-1}$ are $n - 1$ arbitrary functions, bounded and measurable in S.

The solution of $\hat{I}^{(n)}$ depends on the choice of the ϕ_j. For this reason a minimizing function in this problem will be denoted by $w^{(n)}[\phi_j]$ or more briefly by $\hat{w}^{(n)}$. For the corresponding minimum we shall use the symbol $\lambda^{(n)}[\phi_j]$ or $\hat{\lambda}^{(n)}$.

In order to define the intermediate problems giving lower bounds for $\lambda^{(n)}$ we introduce, as in *W*, an arbitrarily given sequence

$$p_1(x,y),\, p_2(x,y),\cdots \tag{1.5}$$

of harmonic functions of integrable square in S.[3] Let us consider the following problem.

PROBLEM $I_m^{(n)}$. *To find a minimizing function for the expression $I(w)/H(w)$ among all functions w continuous in $S+C$, possessing in S continuous first and second derivatives and a Laplacian Δw of integrable square in S. Moreover w must satisfy the following conditions*

$$w = 0 \qquad \text{on } C, \tag{1.6}$$

$$(p_k, \Delta w) = 0 \qquad \text{for } k = 1, 2, \cdots, m. \tag{1.7}$$

The m conditions (1.7) express the orthogonality of the Laplacian Δw with respect to the linear manifold $\mathfrak{P}_m$ of harmonic functions spanned by the m functions $p_1, \cdots, p_m$.

For $n > 1$, w is required to satisfy the $n-1$ additional orthogonality conditions

$$(w, w_m^{(j)}) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1, \tag{1.8}$$

where the notation $w_m^{(i)}$ is used in a general way for a solution of $I_m^{(i)}$. The minimum in $I_m^{(n)}$ will be denoted by $\lambda_m^{(n)}$.

For $n > 1$ it will be useful to consider a Problem $\hat{I}_m^{(n)}$ generalizing $I_m^{(n)}$. The formulation of $\hat{I}_m^{(n)}$ differs from that of $I_m^{(n)}$ only in respect to the condition (1.8), instead of which the following $n-1$ orthogonality conditions are required:

$$(w, \phi_j) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1, \tag{1.9}$$

where the ϕ_j denote, as in $\hat{I}^{(n)}$, arbitrary functions bounded and measurable in S. We denote by $w_m^{(n)}[\phi_j]$ a solution of $\hat{I}_m^{(n)}$ and by $\lambda_m^{(n)}[\phi_j]$ the minimum in this problem. The abbreviated notations $\hat{w}_m^{(n)}$ and $\hat{\lambda}_m^{(n)}$ will sometimes be used.

---

[3] We consider in the present paper only harmonic functions which are *regular* in S. A harmonic function of integrable square can therefore have singularities only on the boundary of S.

The membrane problem can be considered as a limiting case of the problems $I_m^{(n)}$. In fact, by cancelling the conditions (1.7) or, what amounts to the same, by substituting for (1.7) the condition $(p_0, \Delta w) = 0$, where p_0 is the harmonic function 0, we obtain, for $n = 1, 2, \cdots$, a set of problems $I_0^{(n)}$ giving us the eigen-functions and the squares of the eigenvalues of the membrane (see W. p. 19, footnote 2).

Another limiting case—which will be denoted by $I_\infty^{(n)}$—will play an important part in the present paper.

PROBLEM $I_\infty^{(n)}$. *The formulation of* $I_\infty^{(n)}$ *can be obtained from the formulation of* $I_m^{(n)}$ *by the following changes. The index* m *is to be replaced by the index* ∞ *everywhere except in* (1.7) *which is to be replaced by the condition*

$$(p, \Delta w) = 0 \tag{1.10}$$

expressing the orthogonality of the Laplacian Δw *to any harmonic function* p *of the set* $\mathfrak{P}$ *of all harmonic functions of integrable square in* S.[4]

Here again we shall have to consider a Problem $\hat{I}_\infty^{(n)}$ generalizing $I_\infty^{(n)}$ in the same way as $\hat{I}_m^{(n)}$ generalizes $I_m^{(n)}$.

The following inequalities are obviously true for any given set of arbitrary functions ϕ_j:

$$\lambda_m^{(n)}[\phi_j] \leqq \lambda_{m+1}^{(n)}[\phi_j] \leqq \lambda_\infty^{(n)}[\phi_j]. \tag{1.11}$$

It is known from the general theory of eigenvalues that the maximum of the minima in $\hat{I}^{(n)}$, obtained by varying the ϕ_j, is equal to the minimum in $I^{(n)}$ and that this maximum minimorum can be obtained by taking for the ϕ_j the solutions $w^{(j)}$ of the problems $I^{(j)}$ $(j = 1, 2, \cdots, n-1)$. For this choice of the ϕ_j the problems $\hat{I}^{(n)}$ and $I^{(n)}$ become identical. A similar statement is true for the problems $\hat{I}_m^{(n)}$ and $\hat{I}_\infty^{(n)}$ disclosing their connection with $I_m^{(n)}$ and $I_\infty^{(n)}$. As an immediate consequence of these well-known facts we can deduce from (1.11) the inequalities

$$\lambda_m^{(n)} \leqq \lambda_{m+1}^{(n)} \leqq \lambda_\infty^{(n)}. \tag{1.12}$$

In **3** we shall prove the existence of the solutions in $I_m^{(n)}$ and $I_\infty^{(n)}$ (as well as in the important auxiliary problems $\hat{I}_m^{(n)}$ and $\hat{I}_\infty^{(n)}$) for any bounded domain S of finite connectivity under the assumption that none of the boundary components of S consists of a single point.

---

[4] It is known that a strong limit (see **2**) of harmonic functions of the Hilbert space of all harmonic functions of integrable square in S is itself a harmonic function. Hence the set $\mathfrak{P}$ is identical with the closed linear manifold spanned by all harmonic functions of integrable square in S. Cf. S. Zaremba, "Sur l'intégration de l'équation biharmonique," *Bulletin de l'Académie de Cracovie* (1908), pp. 1-29.

In **4** we prove that $\lambda_m^{(n)}$ tends to $\lambda_\infty^{(n)}$ in the case where the sequence (1.5) of harmonic functions p_k, of integrable square in S, is complete in this domain with respect to $\mathfrak{P}$. (The terminology used here is the following: A sequence $\{\psi_n\}$ is said to be complete with respect to a certain class K of functions if the conditions $(\psi, \psi_n) = 0$ imply $\psi = 0$ for every ψ of K.)

In **5** we shall see that, for a very general class $\mathfrak{S}_1$ of domains, the limiting problem $I_\infty^{(n)}$ is equivalent to the classical problem $I^{(n)}$.[5] In this case we have $\lambda_\infty^{(n)} = \lambda^{(n)}$ so that $\lambda_1^{(n)}, \lambda_2^{(n)}, \cdots$ is a non-decreasing sequence of lower bounds for the eigenvalue $\lambda^{(n)}$.

The intermediate problems $I_m^{(n)}$ (or $\hat{I}_m^{(n)}$) differ from the modified problems introduced in W (§ 5) only in the assumption that the harmonic functions p_k belong to the more general class of functions of integrable square in S. This assumption will lead us in a natural way to a geometric interpretation of the problems $I_m^{(n)}$, $\hat{I}_m^{(n)}$, $I_\infty^{(n)}$, $\hat{I}_\infty^{(n)}$ in Hilbert space, which will prove useful in several ways.

Remark. As may be seen in W, § 14, a special sequence of harmonic functions, called " suite privilegiée régulière," has been used for the integration of the differential problem corresponding to $I_m^{(n)}$. It has been recently proved that this sequence is complete in the space $\mathfrak{P}$.[6] From this follows that the convergence theorem remains valid if (1.5) is the " suite privilegiée régulière."

In **6** we prove for the domains of the class $\mathfrak{S}_1$ that the variational problem $I^{(n)}$ is equivalent to the differential problem of the vibrating clamped plate. Using the results of **5** we see at once that an equivalent differential problem is obtained if instead of the conditions (1.1), (1.2) we postulate the conditions (1.6) and (1.10). Let us note that as far back as 1909 Zaremba introduced these last conditions in the theory of the bending of a clamped plate.[7] However he did not investigate the connection between his problem and other problems of a more elementary nature. The reader may be reminded that the differential problem corresponding to $I_m^{(n)}$ differs from the one corresponding to $I_\infty^{(n)}$ not only in the substitution of the orthogonality conditions (1.7) for the condition (1.10), but also by an additional *natural boundary condition.* This condition plays an essential part in the integration of the intermediate problems in terms of the solutions of the membrane problem (see W, § 6 and § 7, as well as the footnote 9 of the present paper.)

---

[5] The existence theorems for the problems $I^{(n)}$ considered in a domain of the class $\mathfrak{S}_1$ is an immediate consequence of this equivalence.

[6] N. Aronszajn, " Sur une propriété de la fonction de Green et sur une suite de fonctions harmoniques introduites par A. Weinstein." This paper was intended to be published in the *Bulletin de la Société Mathématique de France*. Let us mention that it has been already shown in W § 14 that the " suite privilegiée " is complete on the boundary C of S with respect to the class of functions continuous on C.

The notations $\lambda^{(n)}$, $\lambda_m^{(n)}$, $w^{(n)}$, $w_m^{(n)}$ used in the present paper correspond to the notations $\lambda^2{}_n$; $\mu^2{}_{nm}$, w_n and w_{nm} in W.

[7] S. Zaremba, " Sur le problème biharmonique restreint," *Annales de l'Ecole Normale supérieure*, 26 (1909), p. 337.

2. Summary of some facts concerning the Green's potential. In the following a functional operator will be frequently considered. Its definition involves the notion of the Green's function. We assume that the Green's function g exists for our bounded domain S in the (x,y)-plane. It is well known that the existence of g is proved for domains of finite connectivity provided that the boundary of the domain does not contain isolated points.

Let $f(x,y)$ be a function of integrable square in S. Let us consider the operator

$$Gf = G(x,y;f) = -\int\int_S g(x,y,\xi,\eta) f(\xi,\eta)\, d\xi d\eta.$$

Gf, considered as a function of x, y, is usually called the Green's potential of f. It is well known that $G(x,y;f)$ is a continuous function of (x,y) in $S+C$, vanishing on C. If f is bounded, G possesses continuous first derivatives in S. If the first derivatives of f are bounded in S, G possesses continuous second derivatives in S.

Let us now consider Gf as a linear operator in the space $\mathfrak{H}$ of functions f of integrable square in S. Gf defines a transformation of $\mathfrak{H}$ into itself, i. e. Gf is, for an f of integrable square in S, also a function of integrable square. It is easy to see that Gf is a completely continuous operator, i. e. that if a sequence $\{f_k\}$ converges weakly to f, the sequence $\{Gf_k\}$ converges in the strong sense to Gf.

Let us recall the definitions used here: A sequence $\{f_k\}$ in $\mathfrak{H}$ is said to converge in the weak sense to f if $\lim (f_k, \phi) = (f, \phi)$ holds for every function ϕ in $\mathfrak{H}$. $\{f_k\}$ is said to converge in the strong sense to f if we have $\lim (f - f_k, f - f_k) = 0$.

Let us also recall the following properties of the scalar product which will be used in this paper.

1. Let $\{f_k\}$ be a sequence weakly converging to f and $\{\phi_k\}$ a sequence strongly converging to ϕ. Then we have $\lim (f_k, \phi_k) = (f, \phi)$.

2. If $\{f_k\}$ is a sequence weakly converging to f, $\liminf (f_k, f_k) \geqq (f, f)$.

The operator G is symmetric, i. e. we have for any two functions f and ϕ the equation $(f, G\phi) = (Gf, \phi)$. Furthermore we have $(Gf, Gf) > 0$ and $(f, Gf) < 0$ for any $f \neq 0$. The second inequality states that G is negative-definite.

Let w be a continuous function possessing first and second derivatives continuous in S as well as a Laplacian $\Delta w = f$ of integrable square in S. It is well-known that the equation $\Delta w = f$ is then equivalent to the equation

$$w = Gf + p \tag{2.1}$$

where p is harmonic. If w is continuous in $S + C$, the function p assumes on C the same values as w. If, in particular, w vanishes on C we have $w = Gf$.

If f is a function of integrable square in S and $w = Gf$, we have $\Delta w = f$ almost everywhere in S.[3]

3. Existence theorems for the problems $I_m^{(n)}$ and $I_\infty^{(n)}$. It will obviously be sufficient to establish the existence theorems for the problems $\hat{I}_m^{(n)}$ and $\hat{I}_\infty^{(n)}$ which are more general than the problems $I_m^{(n)}$ and $I_\infty^{(n)}$ respectively. We begin our investigation by giving a new formulation for these problems in order to make available a geometric interpretation in Hilbert space. To this end we consider the Laplacian $\Delta w = f$ as the new unknown function instead of w. The function f must satisfy the following conditions equivalent to the conditions for w in $\hat{I}_m^{(n)}$ and $\hat{I}_\infty^{(n)}$:

1°. The equation $\Delta w = f$, where w is continuous in $S + C$ and vanishes on C, is equivalent to the equation

$$w = Gf. \tag{3.1}$$

2°. Using (3.1) we may write, instead of the orthogonality conditions (1.4), the equations $(Gf, \phi_j) = 0$. G being symmetric we obtain finally in place of (1.4) the equivalent conditions

$$(f, G\phi_j) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1. \tag{3.2}$$

3°. We have obviously

$$(p_k, f) = 0 \qquad \text{for } k = 1, 2, \cdots, m \tag{3.3}$$

and

$$(p, f) = 0 \tag{3.4}$$

in place of the orthogonality conditions (1.7) and (1.10) of $\hat{I}_m^{(n)}$ and $\hat{I}_\infty^{(n)}$.

---

[3] Cf. L. Lichtenstein, *Journal für reine und angewandte Mathematik*, 141 (1912), p. 12.

4°. Furthermore we have obviously

$$I(w) = (\Delta w, \Delta w) = (f, f) \tag{3.5}$$

and, by (3.1),

$$H(w) = (w, w) = (Gf, Gf). \tag{3.6}$$

With the present notation we have to determine the minimum of

$$(f, f)/(Gf, Gf) \tag{3.7}$$

in place of the minimum of I/H. We denote by $II_m^{(n)}$, $\hat{II}_m^{(n)}$, $II_\infty^{(n)}$ and $\hat{II}_\infty^{(n)}$ the problems corresponding to $I_m^{(n)}$, $\hat{I}_m^{(n)}$, $I_\infty^{(n)}$, $\hat{I}_\infty^{(n)}$. Obviously we can restrict ourselves to the investigation of $\hat{II}_m^{(n)}$ and $\hat{II}_\infty^{(n)}$, the problems $II_m^{(n)}$ and $II_\infty^{(n)}$ being only special cases of the former.

PROBLEM $\hat{II}_m^{(n)}$. *To prove the existence of a minimizing function for the expression* (3.7) *among all non-identically vanishing functions f of the space $\mathfrak{H}$, orthogonal to the harmonic functions of the subspace $\mathfrak{P}_m$. For $n > 1$ our functions f have to satisfy the additional orthogonality conditions $(f, G\phi_j) = 0$ for $j = 1, 2, \cdots, n-1$.* A solution of $\hat{II}_m^{(n)}$ will be denoted by $f_m^{(n)}[\phi_j]$ or more briefly by $\hat{f}_m^{(n)}$. As will be seen a few lines later, the minimum in $\hat{II}_m^{(n)}$ is equal to the minimum $\lambda_m^{(n)}$ in $I_m^{(n)}$. By substituting $\hat{II}_m^{(n)}$ for $\hat{I}_m^{(n)}$ we have obtained a geometrical formulation of our problem in Hilbert space.

It is clear that there exist functions satisfying the conditions in $\hat{II}_m^{(n)}$ and that the lower bound $\hat{\mu}$ of the expression (3.7) is finite for all such functions. In order to prove the existence of $\hat{f}_m^{(n)}$, let us consider a sequence $\{f_k\}$ of admissible functions for which our functional (3.7) converges to its lower bound $\hat{\mu}$. We may obviously assume, without loss of generality, that the functions f_k are normalized, i. e. $(f_k, f_k) = 1$, so that the expression (3.7) becomes, for $f = f_k$, equal to $1/(Gf_k, Gf_k)$. As the norms of the f_k are bounded we can extract from the sequence $\{f_k\}$, according to the general theory of Hilbert space, a weakly convergent subsequence. To avoid unnecessary complications let us assume that the sequence $\{f_k\}$ itself is already convergent in the weak sense and let us denote by $\hat{f}$ its weak limit. Since the operator G is completely continuous, the functions Gf_k converge to $G\hat{f}$ in the strong sense. Using a result quoted in **2** we see that the non-decreasing sequence (Gf_k, Gf_k) converges to $(G\hat{f}, G\hat{f})$ so that we have

$$\hat{\mu} = \lim_{k=\infty} \frac{1}{(Gf_k, Gf_k)} = \frac{1}{(G\hat{f}, G\hat{f})}.$$

The quantities (Gf_k, Gf_k) being positive and non-decreasing, $(G\hat{f}, G\hat{f})$ is obviously positive so that $\hat{f}$ does not vanish identically. $\hat{f}$ satisfies the conditions of our problem; in fact we have

$$(\hat{f}, G\phi_j) = \lim\,(f_k, G\phi_j) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1,$$
$$(\hat{f}, p) = \lim\,(f_k, p) = 0$$

for all p in $\mathfrak{P}_m$. Since $(f_k, f_k) = 1$ we have, according to **2**, $(\hat{f}, \hat{f}) \leqq 1$. Therefore the quantity $(\hat{f}, \hat{f})/(G\hat{f}, G\hat{f})$ is not greater than $1/(G\hat{f}, G\hat{f})$ which is the lower bound of the functional (3.7). This result proves that $\hat{f}$ is a solution $\hat{f}_m^{(n)}$ of the problem $\hat{\Pi}_m^{(n)}$.

Let us now obtain for $\hat{f}$ a functional equation corresponding to the Euler equation of the variational calculus. In order to obtain this equation let us consider an arbitrary function ϕ of $\mathfrak{H}$ orthogonal to the functions $G\phi_j$ as well as to the functions of $\mathfrak{P}_m$. Let $f = \hat{f} + \epsilon\phi$. Then we have

$$\frac{(\hat{f}, \hat{f})}{(G\hat{f}, G\hat{f})} \leqq \frac{(\hat{f} + \epsilon\phi, \hat{f} + \epsilon\phi)}{(G[\hat{f} + \epsilon\phi], G[\hat{f} + \epsilon\phi])},$$

which gives us the inequality

$$(\hat{f}, \hat{f})\,[2\epsilon(G\hat{f}, G\phi) + \epsilon^2(G\phi, G\phi)] \leqq (G\hat{f}, G\hat{f})\,[2\epsilon(\hat{f}, \phi) + \epsilon^2(\phi, \phi)]$$

which is true for all real values of ϵ in the neighborhood of 0. This inequality can hold only if we have

$$(3.8) \qquad (G\hat{f}, G\hat{f})\,(\hat{f}, \phi) = (\hat{f}, \hat{f})\,(G\hat{f}, G\phi)$$

By the symmetry of the operator G we can write $(GG\hat{f}, \phi)$ in place of $(G\hat{f}, G\phi)$ so that we obtain in place of (3.8) the equation

$$(\hat{f}, \phi) = \frac{(\hat{f}, \hat{f})}{(G\hat{f}, G\hat{f})}(GG\hat{f}, \phi) = \hat{\mu}(GG\hat{f}, \phi) = (\hat{\mu}GG\hat{f}, \phi)$$

or, finally, the equation

$$(3.9) \qquad (\hat{f} - \hat{\mu}GG\hat{f}, \phi) = 0,$$

which holds for all functions ϕ orthogonal to the $G\phi_j$ and to the functions of the space $\mathfrak{P}_m$. The set of these functions ϕ is a linear manifold Φ completely orthogonal to the linear manifold spanned by the $G\phi_j$ and $\mathfrak{P}_m$. By (3.9) and the general projection theorem this last manifold contains the function $\hat{f} - \hat{\mu}GG\hat{f}$ so that we may write

$$\hat f = \hat\mu G \hat f + \sum_{j=1}^{n-1} b_j G\phi_j + p \tag{3.10}$$

where p is a harmonic function of $\mathfrak{P}_m$ and $b_1, \cdots, b_{n-1}$ denote $n-1$ constants. Both sides of this equation are functions of $\mathfrak{H}$ and the equality (3.10) holds only almost everywhere in S. However it is obviously allowable to modify the definition of a solution of our variational problem in a set of zero measure so that we can choose a function $\hat f$ satisfying the equation (3.10) everywhere in S. It may be easily seen by the "Euler equation" (3.10) that $\hat f$ is continuous and has continuous derivatives of the first order in S. Therefore, by **2**, the function $\hat w = G\hat f$ is continuous and has continuous first and second derivatives in S. $\hat w$ is obviously a solution $\hat w_m^{(n)}$ of $\hat I_m^{(n)}$ so that $\hat\mu = \hat\lambda_m^{(n)}$.

In the problem $II_m^{(n)}$ the conditions

$$(f, GGf_m^{(j)}) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1, \tag{3.11}$$

stand for the conditions $(f, G\phi_j) = 0$. We have written $f_m^{(j)}$ for $\Delta w_m^{(j)}$ for the time being. As in the former case, each solution of $II_m^{(n)}$ gives us a solution of $I_m^{(n)}$.

The existence of a solution $\bar f$ in $II_m^{(n)}$ can be proved in the same way as in $\hat{II}_m^{(n)}$. It may however be noted that the Laplacians $\Delta w_m^{(i)}$ of the solutions $w_m^{(i)}$ of $I_m^{(i)}$ $(1 \leqq i \leqq n-1)$ appear in the conditions (3.11) of $II_m^{(n)}$, so that the proof of the existence of the solutions of $II_m^{(n)}$ involves an induction process.

Let us denote for a moment the minimum in $II_m^{(n)}$ by $\bar\mu$. We obtain then, in place of (3.9), the equation

$$(\bar f - \bar\mu GG\bar f, \phi) = 0 \tag{3.12}$$

where ϕ stands for any function orthogonal to the functions $GGf_m^{(j)}$ $(j = 1, 2, \cdots, n-1)$ and to the functions of $\mathfrak{P}_m$.

We deduce from (3.12) an equation similar to (3.10):

$$\bar f = \bar\mu GG\bar f + \sum_{j=1}^{n-1} b_j GGf_m^{(j)} + p \tag{3.13}$$

and as in $\hat{II}_m^{(n)}$ we obtain the result that $\bar\mu$ is identical with the minimum $\lambda_m^{(n)}$ in $I_m^{(n)}$ and that the function $G\bar f$ is a solution $w_m^{(n)}$ of this last problem. Therefore we can write everywhere $\lambda_m^{(n)}$ instead of $\bar\mu$ and $f_m^{(n)}$ in place of $\bar f$.

Putting in (3.12) $\phi = f_m^{(i)}$ where $i > n$ and using (3.11) for $f = f_m^{(n)}$ we obtain the orthogonality equations

$$(f_m^{(n)}, f_m^{(i)}) = 0. \tag{3.14}$$

Exchanging the indices i and n we see that (3.14) holds for any values of $i \neq n$. Let us now multiply (3.13) scalarly by $f_m^{(i)}$ $(i = 1, 2, \cdots, n-1)$. Using (3.11) and (3.14), where we put $f_m^{(n)}$ in place of f, we obtain $b_j = 0$ for $j = 1, 2, \cdots, n-1$, so that we have, instead of (3.13), the following "Euler equation":

$$f_m^{(n)} = \lambda_m^{(n)} GGf_m^{(n)} + p \tag{3.15}$$

where p denotes a function of $\mathfrak{P}_m$.[9]

Problems $II_\infty^{(n)}$ *and* $\hat{II}_\infty^{(n)}$. The formulation of these problems is obtained by replacing in $II_m^{(n)}$ and $\hat{II}_m^{(n)}$ the index m by the index ∞ and the manifold $\mathfrak{P}_m$ by the manifold $\mathfrak{P}$ of all harmonic functions of integrable square in S. The existence theorems can be obtained in the same way as before. As the existence of admissible functions in $II_\infty^{(n)}$ or $\hat{II}_\infty^{(n)}$ may seem at a first glance less obvious than in the cases of $II_m^{(n)}$ or of $\hat{II}_m^{(n)}$ let us mention that the Laplacian $f = \Delta w$ of any function w (continuous and having second derivatives in S and moreover vanishing identically in the neighborhood of the boundary C) is orthogonal to $\mathfrak{P}$.

In the present case we have in place of (3.15) the equation

$$f_\infty^{(n)} = \lambda_\infty^{(n)} GGf_\infty^{(n)} + p \tag{3.16}$$

where p is a function of $\mathfrak{P}$. Let us also note that

$$\lim_{n\to\infty} \lambda_\infty^{(n)} = \infty. \tag{3.17}$$

In fact, $f_\infty^{(n)}$ being not identically zero, we may assume that $(f_\infty^{(n)}, f_\infty^{(n)}) = 1$. On the other hand, putting in (3.14) the index ∞ instead of the index m, we have $(f_\infty^{(n)}, f_\infty^{(i)}) = 0$ for $i \neq n$. Since an orthogonal normalized sequence always converges weakly to zero,[10] we have $\lim\, (Gf_\infty^{(n)}, Gf_\infty^{(n)}) = (G0, G0) = 0$ and

$$\lim_{n\to\infty} \lambda_\infty^{(n)} = \lim_{n\to\infty} \frac{1}{(Gf_\infty^{(n)}, Gf_\infty^{(n)})} = \infty.$$

4. Convergence theorems. From now on we make the fundamental assumption that the sequence of harmonic functions (1. 5), until now arbitrary, is complete in the space $\mathfrak{P}$ [11] and we prove the following theorem.

---

[9] We have $p = a_1p_1 + \cdots + a_mp_m$ where $a_1, \cdots, a_m$ are constants. Assuming that the functions $p_1, \cdots, p_m$ are continuous in $S + C$ we obtain from (3.15) immediately the following *natural boundary condition*: $f_m^{(n)} = \Delta w_m^{(n)} = a_1p_1 + \cdots + a_mp_m$ on C. In the same way we see that the differentiability of the $p_1, \cdots p_m$ in $S + C$ implies the differentiability of $\Delta w_m^{(n)}$ on regular boundaries C. These facts were used in *W*.

[10] This fact is an obvious consequence of the inequality of Bessel.

CONVERGENCE THEOREM. *Assuming that the sequence* (1. 5) *is complete in* $\mathfrak{P}$ *we have*

$$\lim_{m\to\infty} \lambda_m{}^{(n)} = \lambda_\infty{}^{(n)} \qquad \textit{for } n = 1, 2, \cdots \tag{4.1}$$

Let us prove first that the minimum $\lambda_m{}^{(n)}[\phi_j]$ in $\hat{I}_m{}^{(n)}$ tends to the minimum $\lambda_\infty{}^{(n)}[\phi_j]$ in $\hat{I}_\infty{}^{(n)}$ for any fixed value of n and for m converging to infinity. Assuming the set ϕ_j and the index n to be fixed, we use the abbreviations

$$f_m{}^{(n)}[\phi_j] = \hat{f}_m \text{ and } \lambda_m{}^{(n)}[\phi_j] = \hat{\lambda}_m \qquad (m = 1, 2, \cdots \text{ and } m = \infty), \tag{4.2}$$

so that the equation to be proved may be written

$$\lim_{m\to\infty} \hat{\lambda}_m = \hat{\lambda}_\infty. \tag{4.3}$$

Without loss of generality we may assume that $(\hat{f}_m, \hat{f}_m) = 1$. There exists, therefore, a subsequence $\{\hat{f}_{m_k}\}$ of the sequence $\{\hat{f}_m\}$ converging weakly to a function f^*. Obviously f^* is orthogonal to the functions $G\phi_j$ as well as to the functions of the sequence (1. 5). Since the sequence (1. 5) is complete in $\mathfrak{P}$ the function f^* is orthogonal to any function p of $\mathfrak{P}$ as required in problem $\widehat{II}_\infty{}^{(n)}$.

Let us consider now the limit of the non-decreasing sequence $\hat{\lambda}_1 \leqq \hat{\lambda}_2 \leqq \cdots$. This limit is equal to the limit of the subsequence $\{\hat{\lambda}_{m_k}\}$. The condition of $\widehat{II}_\infty{}^{(n)}$ being more stringent than those of $\widehat{II}_m{}^{(n)}$ $(m = 1, 2, \cdots)$, we have

$$\lim \hat{\lambda}_m \leqq \hat{\lambda}_\infty. \tag{4.4}$$

On the other hand we have

$$\lim \hat{\lambda}_m = \lim \hat{\lambda}_{m_k} = \lim \frac{1}{(G\hat{f}_{m_k}, G\hat{f}_{m_k})} = \frac{1}{(Gf^*, Gf^*)}. \tag{4.5}$$

The equations (4. 4) and (4. 5) give us the following result

$$\lim \hat{\lambda}_m = \frac{1}{(Gf^*, Gf^*)} \leqq \hat{\lambda}_\infty. \tag{4.6}$$

Since $\hat{\lambda}_\infty$ is finite, we have $(Gf^*, Gf^*) > 0$ so that f^* does not vanish identically. For this reason f^* is an admissible function in the problem $\widehat{II}_\infty{}^{(n)}$

---

[11] The orthogonality of the p_k is not required in our proof. It is important to notice this fact since the "suite privilegiée régulière" used in W is not always orthogonal.

and we have therefore $(f^*, f^*)/(Gf^*, Gf^*) \geqq \hat{\lambda}_\infty$. Since $\{\hat{f}_{m_k}\}$ converges weakly to f^*, we have, by a result quoted in **2**, $(f^*, f^*) \leqq \liminf (\hat{f}_{m_k}, \hat{f}_{m_k}) = 1$, so that we have

$$(4.7) \qquad \lim \hat{\lambda}_m = \frac{1}{(Gf^*, Gf^*)} \geqq \frac{(f^*, f^*)}{(Gf^*, Gf^*)} \geqq \hat{\lambda}_\infty .$$

Comparing (4.6) with (4.7) we see that $\lim \hat{\lambda}_m = \hat{\lambda}_\infty$, or using again the notations of the Introduction, that

$$(4.8) \qquad \lim_{m\to\infty} \lambda_m^{(n)}[\phi_j] = \lambda_\infty^{(n)}[\phi_j].$$

From this result the equation (4.1) can be easily obtained in the following way.

We have from (1.12)

$$(4.9) \qquad \lim_{m\to\infty} \lambda_m^{(n)} \leqq \lambda_\infty^{(n)}.$$

On the other hand we have by the general theory of eigenvalues (see the Introduction) $\lambda_\infty^{(n)}[w_\infty^{(j)}] = \lambda_\infty^{(n)}$. Putting in (4.8) $\phi_j = w_\infty^{(j)}$ we have $\lim\limits_{m\to\infty} \lambda_m^{(n)}[w_\infty^{(j)}] = \lambda_\infty^{(n)}[w_\infty^{(j)}]$ and therefore

$$(4.10) \qquad \lim_{m\to\infty} \lambda_m^{(n)}[w_\infty^{(j)}] = \lambda_\infty^{(n)}.$$

As was already mentioned in the Introduction, the maximum $\lambda_m^{(n)}$ of $\lambda_m^{(n)}[\phi_j]$ is obtained by putting $\phi_j = w_m^{(j)}$ so that we have

$$\lambda_m^{(n)}[w_\infty^{(j)}] \leqq \lambda_m^{(n)}[w_m^{(j)}] = \lambda_m^{(n)}$$

and therefore

$$(4.11) \qquad \lim_{m\to\infty} \lambda_m^{(n)} \geqq \lambda_\infty^{(n)}.$$

Comparing this inequality with (4.9) we obtain immediately the equation (4.1): $\lim\limits_{m\to\infty} \lambda_m^{(n)} = \lambda_\infty^{(n)}$ which is the convergence theorem.

Remark. The convergence theorem remains true if instead of the previous assumption about the sequence (1.5) we postulate that the sequence of the boundary values of the functions $p_k(x, y)$ is complete in the class of all functions of integrable square on the boundary C, provided that the domain S is of the class $\mathfrak{S}_0$ which will be defined in **5**. The proof of this new convergence theorem follows immediately from the previous result, using the fact that, if S is of the class $\mathfrak{S}_0$, the completeness of the sequence $\{p_k\}$ on C implies the completeness of the sequence of the corresponding harmonic functions $p_k(x, y)$ in S.[12] The corresponding harmonic functions are of course to be defined by their boundary values by means of the Green-Poisson formula.

---

[12] Cf. the paper of N. Aronszajn quoted in footnote 6.

5. Equivalence of the problems $I^{(n)}$ and $I_\infty{}^{(n)}$. Both problems are well-defined for any domain S. However, in order to prove the equivalence of $I^{(n)}$ and $I_\infty{}^{(n)}$, we shall have to make the assumptions that the boundary C of S consists of a finite number of continua without isolated points and that the following property is true for S: *To every harmonic function p of integrable square in S and to every positive ϵ there exists a harmonic function $\bar{p}$ continuous in $S+C$ and admitting in S bounded first derivatives so that* $(p-\bar{p}, p-\bar{p}) \leqq \epsilon$. We denote by $\mathfrak{S}_1$ the class of the domains enjoying this property. All domains considered in this paragraph will be of the class $\mathfrak{S}_1$ unless the contrary is explicitly stated. Let us mention that $\mathfrak{S}_1$ contains the class $\mathfrak{S}_0$ of domains bounded by a finite number of simple closed curves, the mean curvature of these curves being bounded except for a finite number of angular points, the angles being > 0 and $< 2\pi$.

The proof of the equivalence of $I^{(n)}$ and $I_\infty{}^{(n)}$ will be given for any value of n. In the case $n > 1$ we assume that the equivalence holds for the indices $1, 2, \cdots, n-1$ and that *one and the same solution* has been chosen for both problems. Our proof will involve therefore an inductive process.

We shall prove first that for any admissible function the boundary conditions

$$(5.1) \qquad w = 0, \qquad \lim w_x = \lim w_y = 0$$

of $I^{(n)}$ imply the conditions

$$(5.2) \qquad w = 0, \qquad (p, \Delta w) = 0 \qquad \text{for } p \text{ in } \mathfrak{P}$$

of $I_\infty{}^{(n)}$. Secondly we shall prove that a *solution* of $I_\infty{}^{(n)}$ satisfies the boundary conditions of $I^{(n)}$. Using these results we shall prove that every solution of $I_\infty{}^{(n)}$ is a solution of $I^{(n)}$, and vice versa, so that $I_\infty{}^{(n)}$ is equivalent to $I^{(n)}$, always provided that the domain S is of the class $\mathfrak{S}_1$.

1°. In order to prove that the conditions (5.1) of $I^{(n)}$ imply the conditions (5.2) of $I_\infty{}^{(n)}$, let us consider a harmonic function $\bar{p}$ satisfying the regularity conditions given in the definition of the class $\mathfrak{S}_1$. Denoting by S_1 a domain completely interior to S and by C_1 its boundary, we have by Green's formula

$$\int\!\!\int_{S_1} \bar{p}\Delta w \, dxdy = -\int_{C_1}\left(\bar{p}\frac{dw}{dn} - w\frac{d\bar{p}}{dn}\right) ds.$$

Assuming that the domain S_1 tends to S in such a way that the length of its boundary C_1 remains bounded (this assumption being permitted since C is rectifiable), we have by the conditions of $I^{(n)}$

$$(5.3) \qquad \int\!\!\int_S \bar{p}\Delta w \, dxdy \equiv (\bar{p}, \Delta w) = 0.$$

Let us consider now a harmonic function p of integrable square in S and let ϵ be an arbitrary positive number. Then there exists a $\bar{p}$ so that $(p-\bar{p}, p-\bar{p}) < \epsilon$. We have $(p, \Delta w) = (\bar{p}, \Delta w) + (p-\bar{p}, \Delta w)$, so that, by (5.3), we have $(p, \Delta w) = (p-\bar{p}, \Delta w) \leqq \sqrt{(p-\bar{p}, p-\bar{p})(\Delta w, \Delta w)} < \sqrt{\epsilon}\sqrt{(\Delta w, \Delta w)}$. Hence for ϵ converging to 0 we obtain the required equation $(p, \Delta w) = 0$. This result can be formulated in the following way: The class of admissible functions in $I_\infty^{(n)}$ contains the class of admissible functions in $I^{(n)}$.

2°. Using the abbreviations $w = w_\infty^{(n)}$, $\lambda = \lambda_\infty^{(n)}$ and putting $\Delta w = f$, we have in place of (3.16) the equation

$$f = \lambda GGf + p. \tag{5.4}$$

Our aim is to prove that w satisfies the conditions (5.1) of $I^{(n)}$. In this proof we shall use an existence theorem for biharmonic functions (Dirichlet's Problem for the equilibrium of plates).[13]

Let us consider a solution of the equation

$$\Delta\Delta v = \lambda w \tag{5.5}$$

where v is the unknown function satisfying the boundary conditions

$$v = 0 \text{ and } \lim v_x = \lim v_y = 0 \tag{5.6}$$

for (x, y) tending to the boundary. It is known that there exists a solution v of (5.5), (5.6) possessing a Laplacian $f_1 = \Delta v$ of integrable square.[14] Since v has continuous second derivatives in S we may write $v = Gf_1$. By the results obtained above, the conditions (5.6) imply the orthogonality of f_1 to all harmonic functions of integrable square.[15]

Using (2.1) we may write instead of $\Delta\Delta v = \Delta f_1 = \lambda w$ the equation $f_1 = \lambda Gw + p_1 = \lambda GGf + p_1$, i.e. the equation

$$\lambda GGf = f_1 - p_1, \tag{5.7}$$

where p_1 is a function of $\mathfrak{P}$. On the other hand we have by (5.4),

$$\lambda GGf = f - p. \tag{5.8}$$

---

[13] R. Courant und D. Hilbert, *Methoden der Mathematischen Physik*, vol. II, Berlin, 1937, Chap. VII, 9, no. 6. This book will be referred to as CH II.

[14] The following result has been proved in CH II, Chap. VII, § 9, no. 6 (In previous existence proofs quoted there somewhat more restrictive assumptions on the boundary were made): a) there exists a solution v of (5.5) of the class $\dot{\mathfrak{K}}$ (defined *loc. cit.*) admitting continuous derivatives v_{xx}, v_{yy}, v_{xy} of integrable square in S; b) this solution v satisfies the conditions (5.6).

[15] This property holds even for the more general domains considered in CH II and is a simple consequence of the fact that v is of the class $\dot{\mathfrak{K}}$ (cf. 7 of the present paper).

Since the decomposition of $\lambda G G f$ into the sum of a harmonic function and of a function orthogonal to all harmonic functions is unique in $\mathfrak{H}$, we have, comparing (5. 7) and (5. 8), the equations $f = f_1$ and $p = p_1$ so that we deduce from $w = Gf$ and $v = Gf_1$ that $w = v$. We see that *a solution* w *of* $I_\infty^{(n)}$ *satisfies the conditions* (5. 1) *of the problem* $I^{(n)}$.[16]

Our proof uses an existence theorem which, in the generality here needed, was published in CH II at the end of 1937. By using some older papers of Zaremba,[17] a second proof of the same theorem can be given for a class of domains more restricted than $\mathfrak{S}_1$ but still larger than $\mathfrak{S}_0$.

The following proof is based on a lemma on the logarithmic potential which can be easily deduced from the considerations of Zaremba. Let us denote by X a point (x, y) of S and let a be the distance of X to the boundary C of S. We shall consider a function $f(X)$ of integrable square in S satisfying uniformly the conditions

$$\lim af(X) = 0 \tag{5.9}$$

for X converging in S to C.

Let $\{X_n\}$ and $\{X'_n\}$ be two sequences of points in the (x, y)-plane with positive distances a_n and a'_n from C, satisfying the conditions

$$a'_n > k_1 a_n, \qquad \overline{X_n X'_n} < k_2 a_n, \qquad \lim a_n = 0 \tag{5.10}$$

where k_1 and k_2 are independent of n. ($\overline{XX'}$ denotes here the distance from X to X'.) It can easily be shown, under these assumptions, that the logarithmic potential

$$v(X) = v(x, y) = \int\int_S \log \sqrt{(x-\xi)^2 + (y-\eta^2)}\, f(\xi, \eta)\, d\xi\, d\eta \tag{5.11}$$

satisfies the following equations

$$\lim_{n=\infty} \left[\frac{\partial v(x_n, y_n)}{\partial x} - \frac{\partial v(x'_n, y'_n)}{\partial x}\right] = \lim_{n=\infty} \left[\frac{\partial v(x_n, y_n)}{\partial y} - \frac{\partial v(x'_n, y'_n)}{\partial y}\right] = \tag{5.12}$$

Let us consider now a class $\mathfrak{S}_2$ of domains S possessing the following property: To every point X of S corresponds in the exterior of S a point X' for which we have

$$a' > k_1 a, \qquad \overline{XX'} < k_2 a,$$

where a and a' are the distances of X and X' to the boundary C, and k_1 and k_2 denote constants independent of X and X'.[18] For X converging to the boundary of a domain S of the class $\mathfrak{S}_2$ the function v (5.11) satisfies by (5.12) uniformly the equations

---

[16] This result is obviously true for the general domains considered in CH II. However the equivalence of $I^{(n)}$ and $I_\infty^{(n)}$ has been proved here only for the class $\mathfrak{S}_1$.

[17] Cf. S. Zaremba, § 12, *loc. cit.*[4]

[18] This class is slightly more general than that of the paper of Zaremba.

$$\lim \left[\frac{\partial v(x,y)}{\partial x} - \frac{\partial v(x',y')}{\partial x}\right] = \lim \left[\frac{\partial v(x,y)}{\partial y} - \frac{\partial v(x',y')}{\partial y}\right] = 0$$

Moreover, if f is orthogonal to the functions p of $\mathfrak{P}$, the logarithmic potential (5.11) satisfies the obvious equations

$$v(x,y) = 2\pi G(x,y;f) \text{ for } (x,y) \text{ in } S,$$
$$v(x',y') = 0 \text{ for } (x',y') \text{ in the exterior of } S.$$

In this case we have

$$\frac{\partial v(x',y')}{\partial x} = \frac{\partial v(x',y')}{\partial y} = 0$$

so that, for $X = (x,y)$ tending to C, we obtain

$$\text{(5.13)} \qquad \begin{aligned} \lim \frac{\partial G(x,y;f)}{\partial x} &= \frac{1}{2\pi} \lim \frac{\partial v(x,y)}{\partial x} = 0 \\ \lim \frac{\partial G(x,y;f)}{\partial y} &= \frac{1}{2\pi} \lim \frac{\partial v(x,y)}{\partial y} = 0. \end{aligned}$$

Let us consider now a solution $f_\infty^{(n)}$ of $II_\infty^{(n)}$. $f_\infty^{(n)}$ is orthogonal to the functions p of $\mathfrak{P}$. Moreover $f_\infty^{(n)}$ satisfies the condition (5.9). In fact we have $f_\infty^{(n)} = \lambda GGf_\infty^{(n)} + p$ and both terms on the right hand side satisfy (5.9). This is obvious for the continuous function $\lambda GGf_\infty^{(n)}$. As for p, it was proved by Zaremba that (5.9) holds for any p of $\mathfrak{P}$. It follows now at once that $Gf_\infty^{(n)}$ satisfies (5.13), so that the solution $w_\infty^{(n)}$ of $I_\infty^{(n)}$ satisfies the conditions of the problem $I^{(n)}$.

3°. Let us consider again a domain S of the class $\mathfrak{S}_1$. Since the conditions of $I^{(n)}$ imply the conditions of $I_\infty^{(n)}$, we have

$$\text{(5.14)} \qquad \lambda_\infty^{(n)} \leqq \lambda^{(n)}.$$

Consider now a solution $w_\infty^{(n)}$ of $I_\infty^{(n)}$. From the previous results we know that $w_\infty^{(n)}$ satisfies the conditions of $I^{(n)}$, so that we have

$$\text{(5.15)} \qquad \lambda_\infty^{(n)} = I(w_\infty^{(n)})/H(w_\infty^{(n)}) \geqq \lambda^{(n)}.$$

Comparing (5.14) with (5.15) we see that $\lambda_\infty^{(n)} = \lambda^{(n)}$, so that every solution $w_\infty^{(n)}$ of $I_\infty^{(n)}$ is also a solution $w^{(n)}$ of $I^{(n)}$. (Incidentally we have thus proved the existence of $w^{(n)}$.) On the other hand a solution $w^{(n)}$ of $I^{(n)}$ satisfies the conditions of $I_\infty^{(n)}$ and gives the same minimum $\lambda^{(n)} = \lambda_\infty^{(n)}$. It follows that *the solutions of $I^{(n)}$ and $I_\infty^{(n)}$ are identical.* (It may be noted that by using the proof based on the results of Zaremba, we obtain the equivalence only for the domains of the class $\mathfrak{S}_1 \cdot \mathfrak{S}_2$.) It is of interest to remark that it has been shown by one of the authors[19] that any domain with either of the following geometrical properties belongs to the class $\mathfrak{S}_1$.

$\mathfrak{A}$. All domains of finite connectivity with a rectifiable boundary consisting of a finite number of arcs of bounded mean curvature, two arcs having at most their end points in common, the angles at the common vertices being > 0 and $\leqq 2\pi$. The angle 0 is admitted in some cases but the general case of the angle 0 has not yet been investigated.

$\mathfrak{B}$. All domains with rectifiable boundary which may be called star domains in the strict sense, i. e. domains where each ray from a certain fixed centre has a single point of intersection with the boundary C of S.

6. Equivalence of the variational problem with the differential problem. Using the results of **3** and **5** we shall prove this equivalence for the domains of the class $\mathfrak{S}_1$. Let us prove first that *every solution* $w^{(n)}$ *of the variational problem* $I^{(n)}$ *is an eigenfunction of the differential equation*

$$\Delta\Delta w = \lambda w \tag{6.1}$$

satisfying the boundary conditions

$$w = 0, \qquad \lim w_x = \lim w_y = 0 \tag{6.2}$$

and corresponding to the eigenvalue $\lambda = \lambda^{(n)}$.

In order to obtain this result let us use the notations

$$w^{(n)} = Gf^{(n)} \qquad f^{(n)} = \Delta w^{(n)}. \tag{6.3}$$

Since a solution of $I^{(n)}$ is also a solution of $I_\infty{}^{(n)}$, we have for $f^{(n)}$ the equation

$$f^{(n)} = \lambda^{(n)} GGf^{(n)} + p, \tag{6.4}$$

which is, with slightly modified notations, identical with (3.16). In fact we have put $f_\infty{}^{(n)} = f^{(n)}$ and have used the equation $\lambda_\infty{}^{(n)} = \lambda^{(n)}$. From (6.4) we obtain, using (6.3), the equation $\Delta w^{(n)} = \lambda^{(n)} Gw^{(n)} + p$. Taking the Laplacian of both sides, we have, finally, the required equation $\Delta\Delta w^{(n)} = \lambda^{(n)} w^{(n)}$.

---

[19] N. Aronszajn, "Approximation des fonctions harmoniques et quelques problèmes de transformation conforme," *Bulletin de la Société Mathématique de France*, 67 (1940), p. 137.

Let us prove now the converse result: *every solution* w *of the differential eigenvalue problem* (6.1), (6.2) *which is continuous in* $S+C$, *and possesses continuous derivatives of the first and second orders in* S *and a Laplacian of integrable square in* S, *is a solution of the variational problem* $I^{(n)}$ *for a certain value of the index* n. This statement implies the fact that every eigenvalue of the differential problem is equal to one of the minima $\lambda^{(n)}$.

To prove this result let us put $\Delta w = f$, $w = Gf$. By (6.1) we have then the following equation

$$f = \lambda GGf + p' \tag{6.5}$$

similar to (6.4). As we have seen, the conditions (6.2) imply the condition $(p'', f) = 0$ for every p'' in $\mathfrak{P}$.

We find, on multiplying (6.5) scalarly by f, the equation

$$\lambda = (f, f)/(Gf, Gf) = I(w)/H(w) \tag{6.6}$$

so that we have the inequality $\lambda^{(1)} \leqq \lambda$. If $\lambda^{(1)} = \lambda$, our statement is proved. Let us assume, therefore, $\lambda^{(1)} < \lambda$. Since, by (3.17), $\lim_{n\to\infty} \lambda^{(n)} = \lim_{n\to\infty} \lambda_\infty{}^{(n)} = \infty$, there is an n such that

$$\lambda^{(n-1)} < \lambda \leqq \lambda^{(n)}. \tag{6.7}$$

Let us now multiply scalarly (6.5) by $f^{(j)}$ $(= \Delta w^{(j)})$ where $w^{(j)}$ denotes a solution of $I^{(j)}$ $(j = 1, 2, \cdots, n-1)$. Using the equations $(f^{(j)}, p') = 0$ we obtain immediately

$$(f, f^{(j)}) = \lambda(Gf, Gf^{(j)}) \qquad \text{for } j = 1, 2, \cdots, n-1. \tag{6.8}$$

On the other hand we have, putting in (6.4) $n = j$ and multiplying this equation scalarly by f,

$$(f, f^{(j)}) = \lambda^{(j)}(Gf, Gf^{(j)}) \qquad \text{for } j = 1, 2, \cdots, n-1. \tag{6.9}$$

Taking the difference of (6.9) and (6.8) we obtain

$$(\lambda - \lambda^{(j)})(Gf, Gf^{(j)}) = 0,$$

i. e., by (6.7)

$$(Gf, Gf^{(j)}) = (w, w^{(j)}) = 0 \qquad \text{for } j = 1, 2, \cdots, n-1. \tag{6.10}$$

Since, by this equation, w is orthogonal to the solutions $w^{(j)}$ of the problems $I^{(j)}$ $(1 \leqq j \leqq n-1)$, the eigenvalue $\lambda = I(w)/H(w)$ cannot be less than $\lambda^{(n)}$. Comparing this result with (6.7) we obtain the equation $\lambda = \lambda^{(n)}$.

Furthermore, the function w satisfying the conditions of $I^{(n)}$ and giving the minimum $\lambda^{(n)}$ of I/H is obviously a solution $w^{(n)}$ of $I^{(n)}$.

Since $I^{(n)}$ and $I_\infty{}^{(n)}$ are equivalent, we also have the result that the variational problems $I_\infty{}^{(n)}$ are equivalent to the differential problem (6.1), (1.6), (1.10).[20]

7. Comparison of the different kinds of boundary conditions. In the bibliography as well as in the present paper we meet several formulations for the "boundary conditions" in the problem of the clamped plate. It is of interest to investigate whether or not these conditions are equivalent.

We shall compare the following conditions:

A. $w = 0$, $\lim w_x = \lim w_y = 0$, $(\Delta w, \Delta w) < \infty$. These are the conditions in $I^{(n)}$.

B. $w = 0$, $dw/dn = 0$ on C, $(\Delta w, \Delta w) < \infty$. These are the classical conditions for a clamped plate. The condition $dw/dn = 0$ is required only at points of the boundary admitting a normal.

C. $w = 0$, $dw/dn = 0$ on C except in a set of linear measure zero, $(\Delta w, \Delta w) < \infty$.

D. $w = 0$ on C, $(\Delta w, \Delta w) < \infty$, $(p, \Delta w) = 0$ for all harmonic functions p of integrable square in S. These are the conditions in $I_\infty{}^{(n)}$ as well as the conditions introduced by Zaremba in the biharmonic Dirichlet problem.

E. The function w belongs to the class $\mathfrak{K}$ defined in CH II, p. 528, no. 6 (cf. also p. 481 and the reference to other papers there).

We shall prove the equivalence of the variational problems of the vibrating plate no matter which "boundary conditions" A, B, C, D or E are used. We assume here, for simplicity, that the domain S is of the subclass $\mathfrak{A}$ of the class $\mathfrak{S}_1$ introduced at the end of **5**.

In the present paragraph our problems will be denoted by $I_A{}^{(n)}$, $I_B{}^{(n)}$ and so on, according to the boundary conditions used.

It is obvious that the conditions A imply B and that B imply C. On the other hand it has been proved that the conditions C are equivalent to D.[21]

---

[20] In the same way we see that for any value of m the variational problems $I_m{}^{(n)}$ are equivalent to the differential problem defined by the equations (6.1), (1.6), (1.7) and the natural boundary condition mentioned in footnote 9.

[21] The proof will be given in a forthcoming paper of N. Aronszajn, "Recherches sur le potentiel logarithmique."

Furthermore we have seen that a solution of the variational problem $I_D^{(n)}$ satisfies the conditions A. Comparing the minima (by a method similar to that used in **5**) we see that the variational problems with the conditions A, B, C, D are equivalent.

Let us consider now the condition E. We prove first that E implies D. In fact one can easily see that a function w satisfying E belongs to the class $\mathfrak{F}$ and to the class $\dot{\mathfrak{D}}$ introduced in CH II, Chapter VII. Using the results of § 4 of the same chapter, we see that w is continuous in $S + C$ and vanishes on C. We shall prove now that w satisfies the third of the conditions D: Δw is orthogonal to every p in $\mathfrak{P}$. In fact, by the definition of $\dot{\mathfrak{K}}$ there exists a sequence of functions $\dot{w}_n$, of the class $\mathfrak{K}$ defined in CH II, p. 528, no. 6, so that $(p, \Delta w) = \lim (p, \Delta \dot{w}_n)$ and $(p, \Delta \dot{w}_n) = 0$.

Consider now a solution w of $I_D^{(n)}$. In order to see that w satisfies the conditions E we put $f = \Delta w$, so that we have

$$f = \lambda^{(n)} GGf + p. \tag{7.1}$$

Consider on the other hand the equation of equilibrium of our plate

$$\Delta\Delta v = \lambda^{(n)} w \tag{7.2}$$

where v is the unknown function satisfying the conditions E. By CH II, Chapter VII, there is in this problem one and only one solution v, which satisfies *a fortiori* the conditions D. Putting $\Delta v = f_1$ and using (7.2) we have

$$f_1 = \lambda^{(n)} GGf + p_1, \tag{7.3}$$

where p_1 is a function of $\mathfrak{P}$. Comparing (7.1) with (7.3) we obtain as in **5** the equations $p_1 = p$ and $f_1 = f$, so that $v = w$. As required we have the result that w satisfies the conditions E. (For the equivalence of A and E cf. also CH II, p. 530).

In order to obtain the equivalence of $I_D^{(n)}$ with $I_E^{(n)}$ we had to use an induction process with respect to the index n in a similar way to that in **5**.

8. Mixed boundary conditions. The problems with mixed boundary conditions,[22] for instance the problems similar to $I^{(n)}$, are obtained by replacing the conditions (1.2) by the condition $\lim \Delta w = 0$ when (x, y) tends to a part Γ of C. We assume that Γ consists of a finite number of arcs with bounded mean curvature. Let us point out the changes which must be made in the formulations of $I_m^{(n)}$ and $I_\infty^{(n)}$. In $I_m^{(n)}$ we have to introduce, instead of (1.5), a sequence $p_1(x, y)$, $p_2(x, y)$, $\cdots$ of harmonic functions of integrable square in S vanishing on Γ. In $I_\infty^{(n)}$ we have to consider in place of $\mathfrak{P}$ the subspace $\mathfrak{P}^{(0)}$ of $\mathfrak{P}$ of harmonic functions vanishing on Γ. This subspace is also a complete space, i. e. a strong limit of functions in $\mathfrak{P}^{(0)}$ is also a function in $\mathfrak{P}^{(0)}$. In fact, let us map conformally the domain S upon a domain S' located in a (ξ, η) plane and bounded by circles, so that certain arcs Γ' correspond to Γ. Consider now the analytic continuation of our harmonic functions over the arcs Γ' using Schwarz' principle. It can easily be verified that an arc corresponding to an arc interior to Γ is located in a domain independent of $p(\xi, \eta)$ in which all $p(\xi, \eta)$ are of integrable square. Furthermore the integral $\int\int p^2 d\xi d\eta$ taken over this domain is less than $M \int\int_S p^2 dxdy$ where the constant M is independent of p.

---

[22] Cf. W § 18.

It is well known that the strong convergence of harmonic functions implies uniform convergence in the interior of a domain. Considering again the initial domain S we see that the strong limit of a sequence in $\mathfrak{P}^{(0)}$ is itself a harmonic function vanishing on Γ. In this way we reach the conclusion that the results in the case of mixed boundary conditions are similar to those in the case of a clamped plate.

9. **Buckling problem.** Let us consider the problem of the buckling of a plate, assuming, for the sake of brevity, that the boundary is clamped. The new problems are obtained by making the following modifications in the variational problems of the vibrating plate.

In $I^{(n)}$ the integral $H(w)$ has to be replaced by the Dirichlet integral.

$$D(w) = \int\int_S (w_x^2 + w_y^2)\, dxdy,$$

and the orthogonality conditions (1.3) by the conditions

$$D(w, w^{(j)}) = \int\int_S (w_x w_x^{(j)} + w_y w_y^{(j)})\, dxdy = 0 \qquad \text{for } j = 1, 2, \cdots, n-1. \tag{9.1}$$

In $\hat{I}^{(n)}$ the conditions (1.4) have to be replaced by the conditions $D(w, \phi_j) = 0$ where the ϕ_j denote arbitrary functions continuous in $S + C$, vanishing on C and possessing bounded first derivatives in S.

In $I_m^{(n)}$ and $\hat{I}_m^{(n)}$, $H(w)$ has to be replaced by $D(w)$ and the conditions (1.8), (1.9) by $D(w, w_m^{(j)}) = 0$ and $D(w, \phi_j) = 0$. Similar changes are to be made in $I_\infty^{(n)}$ and $\hat{I}_\infty^{(n)}$.

In order to formulate the problems corresponding to $II_m^{(n)}$, $\hat{II}_m^{(n)}$, $II_\infty^{(n)}$ and $\hat{II}_\infty^{(n)}$ we use Green's formula

$$D(w, v) = -(\Delta w, v) \tag{9.2}$$

which holds for any two functions w and v, continuous in $S + C$, v vanishing on C. It is assumed that v admits in S continuous first derivatives and that w admits in the same domain continuous second derivatives. Furthermore we assume the existence of the integrals $(\Delta w, \Delta w)$, $D(w)$ and $D(v)$. Let us note that the condition $(\Delta w, \Delta w) < \infty$ can be replaced by the condition that Δw^2 is summable in S, so that we have a set of validity conditions for (9.2) which is invariant under conformal mapping.[23]

We can now formulate the problems corresponding to the problems of **3**, replacing there the expression $(f,f)/(Gf,Gf)$ by $-(f,f)/(Gf,f)$ and $(f, G\phi_j)$ by (f, ϕ_j). Let us mention that the functional $-(Gf,f)$ replacing the functional $(Gf, Gf) = (GGf, f)$ is also completely continuous and positive definite. For this reason all our proofs hold in the present case, the changes in the formulas being very slight. For instance, we obtain in $II_m^{(n)}$ (or in $II_\infty^{(n)}$) instead of (3.15) the equation

$$f = -\lambda Gf + p. \tag{9.3}$$

Let us also mention that in **5** we have to write the equation

$$\Delta\Delta v = -\lambda\Delta w$$

instead of equation (5.5).

London, England,
University of Toronto.

---

[23] The equation (9.2) has already been established in CH II, p. 484 under somewhat different assumptions.

PRODUCTS AND NORMS OF IDEALS.*

By C. C. MacDuffee.

---

1. Introduction. The early work of Poincaré [1] on the representation of ideals in integral domains of algebraic fields by means of matrices suggests that he may have been led into this investigation by the hope that ideal multiplication could in some way be accomplished through multiplication of matrices. It soon became evident that this cannot be done in any simple way, and no further results in this direction have been obtained.

In the present paper this result is accomplished by means of a multiple correspondence. To an ideal $\mathfrak{a}$ correspond, besides the matrix A of the Poincaré correspondence, h matrices $A_1, A_2, \cdots, A_h$ where h is the class number of the domain. If $\mathfrak{b}$ belongs to the ideal class b (where 1 represents the principal class), and if B, C correspond respectively to the ideal $\mathfrak{b}$ and the ideal product $\mathfrak{a} \times \mathfrak{b}$ by the Poincaré correspondence, then

$$C = A_b B$$

where the multiplication is ordinary matric multiplication.

If $\mathfrak{a}$ and $\mathfrak{b}$ are two-sided ideals in a maximal domain of a semi-simple algebra over the rational field, and doubtless also in many other situations, it is true that the norm of the product is equal to the product of the norms of the factors.[2] This is not usually true in the case of the product of one-sided ideals. A satisfactory substitute is given in this paper. Every ideal has, besides its ordinary norm, as many norms as the class number of the domain, and each norm of a product is always equal to the product of certain properly chosen norms of the factors.

2. Notations. The notation in this paper has been chosen to agree with that of an earlier paper of the author,[3] of which this paper is essentially a continuation. We assume that $\mathfrak{F}$ is the quotient field of the principal ideal

---

* Received August 16, 1941; Presented to the American Mathematical Society May 2, 1941.

[1] H. Poincaré, *Bull. Soc. Math. France*, vol. 13 (1885), p. 167.

[2] M. Deuring, *Algebren*, Springer 1935, Chap. VI, § 4.

[3] "Modules and ideals in a Frobenius algebra," *Monatshefte für Math. und Physik*, vol. 48 (1939), pp. 293-313.

ring $\mathfrak{R}$, and that $\mathfrak{A}$ is a finite algebra over $\mathfrak{F}$ with the $\mathfrak{F}$-basis $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$, and the constants of multiplication c_{ijk}, which belong to $\mathfrak{R}$. We assume that $\mathfrak{A}$ is a Frobenius algebra—i. e., that the determinant

$$| l_1 c_{rs1} + l_2 c_{rs2} + \cdots + l_n c_{rsn} |$$

does not vanish for every choice of $l_1, l_2, \cdots, l_n$ in $\mathfrak{F}$. We assume that ϵ_1 is the unit element of $\mathfrak{A}$.

The set $\mathfrak{J}$ of all numbers

$$a_1\epsilon_1 + a_2\epsilon_2 + \cdots + a_n\epsilon_n, \qquad a_i \varepsilon \mathfrak{R}$$

constitutes an integral domain. An additive group $\mathfrak{M}$ of numbers of $\mathfrak{J}$ which is closed under multiplication by the numbers of $\mathfrak{R}$ is called an $\mathfrak{R}$-module. An $\mathfrak{R}$-module which is closed under multiplication on the left by the numbers of $\mathfrak{J}$ is a *left ideal.* A *right ideal* is similarly defined.

Every left or right ideal $\mathfrak{a}$ of $\mathfrak{J}$ has an $\mathfrak{R}$-basis $\alpha_1, \alpha_2, \cdots, \alpha_n$ where

$$\alpha_i = a_{i1}\epsilon_1 + a_{i2}\epsilon_2 + \cdots + a_{in}\epsilon_n, \qquad a_{ij} \varepsilon \mathfrak{R}.$$

Define the matrix $A = (a_{rs})$, and say that A *corresponds* to the ideal $\mathfrak{a}$. If A is non-singular, it is unique up to a left factor U which is a unimodular matrix with elements in $\mathfrak{R}$. The principal left ideal $(\alpha]$ has as a corresponding matrix the matrix $S(\alpha)$ of the second regular representation of $\mathfrak{A}$. Similarly $R^{\mathsf{T}}(\alpha)$ corresponds to the principal right ideal $[\alpha)$, where $R(\alpha)$ is the first regular representation of α.

A necessary and sufficient condition in order that A shall correspond to a left ideal is that there shall exist n matrices $D_1, D_2, \cdots, D_n$ with elements in $\mathfrak{R}$ such that

$$AR_i^{\mathsf{T}} = D_i^{\mathsf{T}}A \qquad R_i = R(\epsilon_i).$$

Similarly A corresponds to a right ideal if and only if matrices $E_1, E_2, \cdots, E_n$ exist such that

$$AS_i = E_iA \qquad S_i = S(\epsilon_i).$$

Let $\mathfrak{a}$ be an $\mathfrak{R}$-module with $\mathfrak{R}$-basis $\alpha_1, \alpha_2, \cdots, \alpha_n$, and $\mathfrak{b}$ an $\mathfrak{R}$-module with basis $\beta_1, \beta_2, \cdots, \beta_n$. The *product* $\mathfrak{a} \times \mathfrak{b}$ is the $\mathfrak{R}$-module

$$\Sigma\, d_{ij}\alpha_i\beta_j, \qquad d_{ij} \varepsilon \mathfrak{R}.$$

If $\mathfrak{a}$ is a left ideal, so is $\mathfrak{a} \times \mathfrak{b}$. If $\mathfrak{b}$ is a right ideal, so is $\mathfrak{a} \times \mathfrak{b}$. If A corresponds to $\mathfrak{a}$, B to $\mathfrak{b}$ and C to $\mathfrak{a} \times \mathfrak{b}$, then C is a greatest common right divisor of the matrices

$$AS(\beta_1), AS(\beta_2), \cdots, AS(\beta_n),$$

and also a g. c. r. d. of the matrices

$$BR^{\mathsf{T}}(\alpha_1), BR^{\mathsf{T}}(\alpha_2), \cdots, BR^{\mathsf{T}}(\alpha_n).$$

If A corresponds to the left ideal $\mathfrak{a}$ with $\mathfrak{R}$-basis $\alpha_1, \alpha_2, \cdots, \alpha_n$, then $AS(\tau)$ corresponds to a left ideal having the $\mathfrak{R}$-basis $\alpha_1\tau, \alpha_2\tau, \cdots, \alpha_n\tau$, called the ideal $\mathfrak{a}\tau$. Two non-singular left ideals $\mathfrak{a}$ and $\mathfrak{b}$ are said to be *equivalent* $(\mathfrak{a} \sim \mathfrak{b})$, or to belong to the same *class*,[4] if there are numbers τ_1 and τ_2 of $\mathfrak{F}$ such that

$$\mathfrak{a}\tau_1 = \mathfrak{b}\tau_2.$$

Let A and B correspond to left ideals $\mathfrak{a}$ and $\mathfrak{b}$, and let

$$AR_i{}^{\mathsf{T}} = D_i{}^{\mathsf{T}}A, \qquad BR_i{}^{\mathsf{T}} = D'_i{}^{\mathsf{T}}B.$$

A necessary and sufficient condition that $\mathfrak{a} \sim \mathfrak{b}$ is that there exist a unimodular matrix U such that

$$D'_i{}^{\mathsf{T}} = U^{-1}D_i{}^{\mathsf{T}}U \qquad (i = 1, 2, \cdots, n).$$

3. Ideal Products. The matrix B corresponding to the ideal $\mathfrak{b}$ is unique only up to a unimodular left factor. If $\mathfrak{b}$ is equivalent to $\mathfrak{a}$ so that $D'_i{}^{\mathsf{T}} = U^{-1}D_i{}^{\mathsf{T}}U$, then

$$BR_i{}^{\mathsf{T}} = U^{-1}D_i{}^{\mathsf{T}}UB, \qquad UBR_i{}^{\mathsf{T}} = D_i{}^{\mathsf{T}}UB,$$

so that UB, which also corresponds to $\mathfrak{b}$, has the same *class matrices* $D_1, D_2, \cdots, D_n$ as $\mathfrak{a}$. We say that two ideal matrices which have the same class matrices belong to the same *minor class.* By a proper choice of the unimodular left factor U, we can determine ideal matrices, all belonging to the same minor class, which correspond to each of the ideals of a class. Every minor class has a *basis* composed of n linearly independent matrices $M_1, M_2, \cdots, M_n$ such that every matrix

$$a_1M_1 + a_2M_2 + \cdots + a_nM_n$$

corresponds to just one ideal of the class of $\mathfrak{a}$. We use the symbol $\sim$ to denote this correspondence. Conversely, every ideal of this class corresponds to at least one such matrix.

Let us enumerate the ideal classes of $\mathfrak{F}$ as the first (or principal), second,

---

[4] If $\mathfrak{F}$ is the rational field, $\mathfrak{R}$ the ring of rational integers, and $\mathfrak{A}$ is semi-simple, the number of classes of left ideals is finite. See Deuring, *loc. cit.*, p. 90.

$\cdots, h$-th, and assume a definite minor class of ideal matrices for every class. We shall denote the class matrices of the i-th minor class by

$$D_1^{(i)}, D_2^{(i)}, \cdots, D_n^{(i)}, \qquad D_j^{(i)} = R_j.$$

Let $\mathfrak{c} = \mathfrak{a} \times \mathfrak{b}$ be the ideal product of the left ideals $\mathfrak{a}$ and $\mathfrak{b}$, and let $\mathfrak{a} \sim A$, $\mathfrak{b} \sim B$, $\mathfrak{c} \sim C$. Then C is a greatest common right divisor of the matrices

$$BR^{\mathsf{T}}(\alpha_1), BR^{\mathsf{T}}(\alpha_2), \cdots, BR^{\mathsf{T}}(\alpha_n).$$

But if B belongs to the b-th minor class,

$$BR^{\mathsf{T}}(\alpha_i) = D^{(b)\mathsf{T}}(\alpha_i)B, \qquad D^{(b)}(\alpha_i) = a_1 D_1^{(b)} + \cdots + a_n D_n^{(b)}$$

so that C is a greatest common right divisor of

$$D^{(b)\mathsf{T}}(\alpha_1)B, D^{(b)\mathsf{T}}(\alpha_2)B, \cdots, D^{(b)\mathsf{T}}(\alpha_n)B.$$

Define A_b to be a g. c. r. d. of

$$D^{(b)\mathsf{T}}(\alpha_1), D^{(b)\mathsf{T}}(\alpha_2), \cdots, D^{(b)\mathsf{T}}(\alpha_n).$$

Then A_bB is a matrix corresponding to the ideal product $\mathfrak{a} \times \mathfrak{b}$. The g. c. r. d. of matrices is unique only up to a unimodular left factor, and so A_b can be chosen so that A_bB lies in the minor class which we have selected to represent the class of this product. That is,

$$C = A_bB$$

where the multiplication is matric multiplication.

Two ideal matrices B and B' belong to the same minor class if and only if a number ν of $\mathfrak{A}$ exists such that $B' = BS(\nu)$ where $S(\nu)$ corresponds to ν under the second regular representation. Now if $C = A_bB$, then

$$CS(\nu) = A_bBS(\nu).$$

If $BS(\nu)$ has elements in $\mathfrak{R}$, so does $CS(\nu)$, and $CS(\nu)$ belongs to the same minor class as C. Thus if we select A_b so that A_bB lies in a prescribed minor class, so does A_bB' for every matrix B' of the b-th minor class.

In this way we have a correspondence wherein every ideal $\mathfrak{a}$ corresponds to h matrices $A_1, A_2, \cdots, A_h$ where h is the class number of $\mathfrak{J}$. This is in addition to the ordinary Poincaré correspondence by which $\mathfrak{a} \sim A$. We have proved

THEOREM 1. *If* $\mathfrak{a} \times \mathfrak{b} = \mathfrak{c}$, *and* $\mathfrak{b}$ *belongs to the b-th minor class, then*

$$C = A_b B$$

corresponds to $\mathfrak{c}$, *and* C *is in the minor class chosen to represent the class of* $\mathfrak{c}$.

In fact, from the matrices corresponding to $\mathfrak{a}$ and $\mathfrak{b}$ we can calculate all the matrices corresponding to the ideal product $\mathfrak{c}$. Let D be any non-singular matrix of the d-th minor class, corresponding to the ideal $\mathfrak{d}$. Since ideal multiplication is associative,

$$\mathfrak{p} = (\mathfrak{a} \times \mathfrak{b}) \times \mathfrak{d} = \mathfrak{a} \times (\mathfrak{b} \times \mathfrak{d}), \qquad \mathfrak{a} \times \mathfrak{b} = \mathfrak{c}, \qquad \mathfrak{b} \times \mathfrak{d} = \mathfrak{e}.$$

Then if P corresponds to $\mathfrak{p}$,

$$P = C_d D = A_c E = A_e B_d D.$$

Since D is non-singular,

$$C_d = A_e B_d$$

where the e-th minor class is the minor class of $B_d D$. As we have seen, e is independent of the particular matrix D chosen from the d-th minor class.

Because of the lack of the commutative law, and the consequent non-existence of the class group, further results seem rather meager.

4. Norms. Let $\mathfrak{a}$ be an ideal of the class a, and let $\mathfrak{a} \sim A$ under the Poincaré correspondence, and let $\mathfrak{a} \sim A_1, A_2, \cdots, A_k$ under the correspondence of Theorem 1. Then $N(\mathfrak{a}) = |A|$ is the ordinary norm of $\mathfrak{a}$. We shall call $N_i(\mathfrak{a}) = |A_i|$ the *i-th norm* of $\mathfrak{a}$.

It is now possible to extend the theorem on the norm of a product to one-sided ideals in the following manner.

THEOREM 2. *If* $\mathfrak{b}$ *is an ideal of the b-th class,*

$$N(\mathfrak{a} \times b) = N_b(\mathfrak{a}) \cdot N(\mathfrak{b}).$$

If D *is any matrix of the d-th minor class and* $B_b D$ *is in the e-th minor class, then*

$$N_d(\mathfrak{a} \times \mathfrak{b}) = N_e(\mathfrak{a}) \cdot N_d(\mathfrak{b}).$$

5. An Example. It was shown by E. J. Finan[5] that the matrices

$$\epsilon_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \epsilon_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \epsilon_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \epsilon_4 = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}$$

---

[5] *Duke Mathematical Journal*, vol. 2 (1935), p. 484.

form a basis of a non-maximal integral ring $\mathfrak{J}$ of the total matric algebra of order 4, that the class number of this ring is 3, and that the three classes can be represented by the three minor classes of matrices

$$A = \begin{pmatrix} a & b & 0 & 0 \\ 3d & c & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & 3b & a \end{pmatrix}, \quad B = \begin{pmatrix} 3a & b & 0 & 0 \\ 3d & c & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & b & a \end{pmatrix}, \quad C = \begin{pmatrix} a & b & 0 & 0 \\ d & c & 0 & 0 \\ 0 & 0 & 3c & d \\ 0 & 0 & 3b & a \end{pmatrix}.$$

We shall call these Classes I, II, III, respectively. Class I is the principal class.

Denote $xD_1^{(i)\mathsf{T}} + yD_2^{(i)\mathsf{T}} + zD_3^{(i)\mathsf{T}} + wD_4^{(i)\mathsf{T}}$ by $D^{(i)\mathsf{T}}$.

Then

$$D^{(1)\mathsf{T}} = \begin{pmatrix} x & 0 & 0 & w \\ 0 & x & 3w & 0 \\ 0 & y & z & 0 \\ 3y & 0 & 0 & z \end{pmatrix}, \quad D^{(2)\mathsf{T}} = \begin{pmatrix} x & 0 & 0 & 3w \\ 0 & x & 3w & 0 \\ 0 & y & z & 0 \\ y & 0 & 0 & z \end{pmatrix}, \quad D^{(3)\mathsf{T}} = \begin{pmatrix} x & 0 & 0 & w \\ 0 & x & w & 0 \\ 0 & 3y & z & 0 \\ 3y & 0 & 0 & z \end{pmatrix}.$$

Let k_1 be the greatest common divisor $(a, 3b, 3c, 3d)$, and similarly let

$$\begin{aligned} k_2 &= (a, b, c, 3d), & k_3 &= (3a, 3b, c, 3d), \\ k_4 &= (a, 3b, c, d), & k_5 &= (3a, b, c, 3d), \\ k_6 &= (a, 3b, 3c, d), & k_7 &= (a, b, c, d). \end{aligned}$$

Denote by $[k_1, k_2, k_3, k_4]$ the diagonal matrix

$$A_1 = \begin{pmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{pmatrix}.$$

Then

$$\begin{aligned} A_2 &= [k_2, k_3, k_3, k_3], & A_3 &= [k_1, k_1, k_4, k_4], \\ B_1 &= [3k_7, k_5, k_5, k_7], & B_2 &= [k_5, k_5, k_5, k_5], \\ B_3 &= [3k_7, 3k_7, k_7, k_7], & C_1 &= [k_6, k_7, 3k_7, k_6], \\ C_2 &= [k_7, k_7, 3k_7, 3k_7], & C_3 &= [k_6, k_6, k_6, k_6]. \end{aligned}$$

Consider the ideal $\mathfrak{a} \sim A$,

$$A_1 A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 1 \end{pmatrix}.$$

Here $\mathfrak{a}$ is in Class I, and so is $\mathfrak{a}^2$. In fact, $\mathfrak{a}$ is idempotent. $N(\mathfrak{a}) = 4$, $N_1(\mathfrak{a}) = 1$, $N(\mathfrak{a}^2) = 4$.

Consider the ideal $\mathfrak{a}' \sim A'$,

$$A'_1A' = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 3 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{pmatrix}.$$

Here $\mathfrak{a}'$ is in Class I but $\mathfrak{a}'^2$ is in Class II. $N(\mathfrak{a}') = 0$, $N_1(\mathfrak{a}') = 3$, $N(\mathfrak{a}'^2) = 0$, so that $\mathfrak{a}'$ is singular.

Consider the ideal $\mathfrak{a}'' \sim A''$,

$$A''_1A'' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 0 & 0 \\ 6 & 6 & 0 & 0 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 6 & 6 & 0 & 0 \\ 0 & 0 & 18 & 6 \\ 0 & 0 & 3 & 2 \end{pmatrix}.$$

Here $\mathfrak{a}''$ is in Class I, $\mathfrak{a}''^2$ is in Class III. $N(\mathfrak{a}'') = 36$, $N_1(\mathfrak{a}'') = 3$, $N(\mathfrak{a}''^2) = 108$. In fact, $N_2(\mathfrak{a}'') = 9$, $N_3(\mathfrak{a}'') = 1$, so all the norms are different.

As we saw in **3**, the class of $\mathfrak{a} \times \mathfrak{b}$ is not altered when $\mathfrak{b}$ is replaced by any other ideal in its class. The above counter-example shows that this is not always true when the left factor is similarly replaced. Thus *left ideal classes do not always form a semi-group under multiplication.*

HUNTER COLLEGE.

NUMERICAL CONTINUED FRACTIONS.*

By J. D. Bankier and Walter Leighton.

This paper is primarily concerned with arithmetic properties of various types of continued fractions. A theorem due to Galois [3, p. 83] is generalized in Section **2**. Later sections deal with properties of a proper continued fraction, a natural generalization of the classical "regular" continued fraction and the continued cotangent of Lehmer [1]. In Section **3**, a result due to Legendre concerning the relations between the complete quotients of regular continued fractions with inverse periods is generalized. In Section **4**, it is shown that a pure periodic proper continued fraction does not, in general, represent a reduced quadratic irrational; sufficient conditions for such a representation are determined, and several theorems of Galois, Legendre and Muir are generalized. A sufficient condition that a certain class of quadratic irrationals will have periodic proper continued fraction expansions is determined in Section **5**, and conditions that a rational number shall be equal to a convergent of a proper continued fraction are given in Section **6**.

1. Conjugate numbers. Convergent pure periodic continued fractions of the form

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_{k-1}}{b_{k-1}} + \frac{a_k}{b_0} + \frac{a_1}{b_1} + \cdots \tag{1.1}$$

will be considered in this section. The elements of (1.1) will be complex numbers with $a_n \neq 0$ $(n = 1, 2, 3, \cdots)$.

The n-th convergent of the continued fraction (1) is denoted by A_n/B_n and the formulas

$$A_{nk-2} = A_{k-1}A_{(n-1)k-2} + a_kA_{k-2}B_{(n-1)k-2}, \tag{1.21}$$

$$A_{nk-2} = A_{(n-1)k-1}A_{k-2} + a_kA_{(n-1)k-2}B_{k-2}, \tag{1.22}$$

$$B_{nk-1} = B_{k-1}A_{(n-1)k-1} + a_kB_{k-2}B_{(n-1)k-1}, \tag{1.23}$$

$$B_{nk-1} = B_{(n-1)k-1}A_{k-1} + a_kB_{(n-1)k-2}B_{k-1}, \tag{1.24}$$

are obtained from the fundamental formulas [3, p. 14]

$$\begin{aligned} A_{\nu+\lambda-1} &= A_{\lambda-1}A_{\nu-1,\lambda} + a_\lambda A_{\lambda-2}B_{\nu-1,\lambda}, \\ B_{\nu+\lambda-1} &= B_{\lambda-1}A_{\nu-1,\lambda} + a_\lambda A_{\lambda-2}B_{\nu-1,\lambda}. \end{aligned} \tag{1.3}$$

---

* Received August 29, 1941.

The value y_0 of the continued fraction (1.1) satisfies the relation [3, p. 16].

$$y_0 = \frac{A_{nk-1}y_0 + a_k A_{nk-2}}{B_{nk-1}y_0 + a_k B_{nk-2}},$$

and hence y_0 is a root y of each equation in the set of equations

$$B_{nk-1}y^2 + (a_k B_{nk-2} - A_{nk-1})y - a_k A_{nk-2} = 0 \qquad (n = 1, 2, 3, \cdots). \tag{1.4}$$

If ν and λ are both set equal to mk in the second of the equations (1.3), it will be found that

$$B_{2mk-1} = B_{mk-1}(A_{mk-1} + a_k B_{mk-2}).$$

Thus, if one of the numbers B_{nk-1} $(n = 1, 2, 3, \cdots)$ were zero, an infinite sequence of these numbers would be zero. But this is impossible since the continued fraction converges. Hence every member of the set of equations (1.4) is a quadratic equation. It will now be shown that these equations all have the same roots. It will be sufficient to prove that

$$\frac{a_k B_{k-2} - A_{k-1}}{B_{k-1}} = \frac{a_k B_{nk-2} - A_{nk-1}}{B_{nk-1}}, \quad \frac{A_{k-2}}{B_{k-1}} = \frac{A_{nk-2}}{B_{nk-1}}, \tag{1.5}$$

for all positive integral values of n. The first of the equalities (1.5) may be established by subtracting equation (1.23) from equation (1.24). The second equality is a consequence of relations (1.22) and (1.23) from which it follows immediately that

$$\begin{aligned} A_{nk-2}B_{k-1} - A_{k-2}B_{nk-1} &= a_k B_{k-2}(A_{(n-1)k-2}B_{k-1} - A_{k-2}B_{(n-1)k-1}) \\ &= (a_k B_{k-2})^{n-1}(A_{k-2}B_{k-1} - A_{k-2}B_{k-1}) = 0. \end{aligned}$$

It is now clear that the equations in the system (1.4) are equivalent to each other. The second root z_0 of these equations will be called the *conjugate* of y_0.

2. A generalization of Galois' theorem.

The generalization of a result due to Galois is contained in the following theorem.

THEOREM 2.1. *If the pure periodic continued fractions*

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_{k-1}}{b_{k-1}} + \frac{a_k}{b_0} + \frac{a_1}{b_1} + \cdots, \tag{2.11}$$

$$\frac{a_k}{b_{k-1}} + \frac{a_{k-1}}{b_{k-2}} + \cdots + \frac{a_2}{b_1} + \frac{a_1}{b_0} + \frac{a_k}{b_{k-1}} + \frac{a_{k-1}}{b_{k-2}} + \cdots, \tag{2.12}$$

converge, and none of the partial numerators a_n is zero, the value of the continued fraction (2.12) is $-z_0$, where z_0 is the conjugate of the value y_0 of the continued fraction (2.11).

Two cases will be considered in the proof of Theorem 2.1, depending on whether or not one of the numbers A_{nk-2} $(n = 1, 2, 3, \cdots)$ is zero. It will first be assumed that none of these numbers is zero. This assumption implies that y_0 is not zero and hence only a finite number of the numbers A_n $(n = 1, 2, 3, \cdots)$ can be zero. It will be convenient in what follows to define $A_{-2} = 0$, $a_0 = 1$, and to make use of the convention that $A_{-1} = 1$, $B_{-1} = 0$, and $B_0 = 1$. Then there exists an integer m $(\geqq -2)$ such that $A_m = 0$, and $A_n \neq 0$ $(n > m)$.

The recursion formulas [3, p. 16]

$$A_n = b_n A_{n-1} + a_n A_{n-2} \qquad\qquad (2.13)$$
$$B_n = b_n B_{n-1} + a_n B_{n-2} \qquad (n = 2, 3, 4, \cdots),$$

may be used to show that

$$a_k \frac{A_{nk-2}}{A_{nk-1}} = \frac{a_k}{b_{k-1}} + \frac{a_{k-1}}{b_{k-2}} + \cdots + \frac{a_{m+3}}{b_{m+2}} = \frac{X_{nk-m-2}}{Y_{nk-m-2}} \qquad (nk - 2 > m),$$

where X_n/Y_n is the n-th convergent of the continued fraction (2.12).

The above expression may be written in the form

$$a_k \frac{A_{nk-2}}{A_{nk-1}} = -y_0 z_0 \frac{B_{nk-1}}{A_{nk-1}} \qquad (nk - 2 > m),$$

and it follows that

$$\lim_{n\to\infty} \left(a_k \frac{A_{nk-2}}{A_{nk-1}} \right) = -z_0.$$

Then, since the continued fraction (2.12) converges, its value must be $-z_0$.

If one of the numbers A_{nk-2} $(n = 1, 2, 3, \cdots)$ is zero, one of the roots y_0, z_0 must be zero, and hence all these numbers are equal to zero. But this implies that y_0 is zero. Furthermore

$$-(y_0 + z_0) = -z_0 = a_k \frac{B_{nk-2}}{B_{nk-1}} - \frac{A_{nk-1}}{B_{nk-1}},$$

and, thus,

$$\lim_{n\to\infty} \left(a_k \frac{B_{nk-2}}{B_{nk-1}} \right) = -z_0.$$

The proof of Theorem 2.1 is now completed by showing, as in the first case, that

$$a_k \frac{B_{nk-2}}{B_{nk-1}} = \frac{X_{nk-m-2}}{Y_{nk-m-2}} \qquad (nk - 2 > m).$$

3. Periodic proper continued fractions. A continued fraction

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots$$

is said to be a proper continued fraction if its elements are integers which satisfy the conditions

$$0 < a_n \leqq b_n \qquad (n = 1, 2, 3, \cdots).$$

The relations

$$b_0 = [y_0], \qquad y_n = \frac{a_n}{y_{n-1} - b_{n-1}}, \qquad b_n = [y_n] \qquad (n = 1, 2, 3, \cdots),$$

may be used to determine the unique proper continued fraction representation of a real number y_0 corresponding to the sequence of positive integers $a_1, a_2, a_3, \cdots$ [2].

If the continued fraction (2.11) is a proper continued fraction, Theorem 2.1 may be expressed in a form more appropriate for the work of this section.

THEOREM 3.1. *If the pure periodic continued fraction*

$$(3.11) \qquad b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_{k-1}}{b_{k-1}} + \frac{a_k}{b_0} + \frac{a_1}{b_1} + \cdots$$

is a proper continued fraction, the continued fraction

$$(3.12) \qquad b_{k-1} + \frac{a_{k-1}}{b_{k-2}} + \frac{a_{k-2}}{b_{k-3}} + \cdots + \frac{a_2}{b_1} + \frac{a_1}{b_0} + \frac{a_k}{b_{k-1}} + \frac{a_{k-1}}{b_{k-2}} + \cdots$$

converges to the value $-a_k/z_0$ *where* z_0 *is the conjugate of the value* y_0 *of the continued fraction* (3.11).

Theorem 3.1 is an immediate consequence of Theorem 2.1. For the continued fraction (3.11) converges, since it is a proper continued fraction, and the continued fraction (3.12) converges since it has only a finite number of distinct elements and these elements are all positive. Division by z_0 is permissible since it is the conjugate of y_0, and y_0 is a quadratic irrational.

Since the value of a periodic proper continued fraction is a quadratic irrational, the complete quotients of the continued fraction (3.11) are all quadratic irrationals of the form

$$y_n = \frac{[D]^{1/2} + P_n}{Q_n} \qquad (n = 0, 1, 2, \cdots),$$

where D is a positive integer which is not a perfect square, and the numbers P_n, Q_n are rational numbers. The complete quotients take on only k distinct values, where k is the number of terms in the primitive period of the continued fraction (3.11). The complete quotients of the continued fractions (3.11) and (3.12) are related to each other in a manner which is described in the next theorem.

THEOREM 3.2. *If the first k complete quotients of the continued fraction* (3.11) *are*

$$\frac{[D]^{1/2}+P_0}{Q_0}, \frac{[D]^{1/2}+P_1}{Q_1}, \ldots, \frac{[D]^{1/2}+P_{k-2}}{Q_{k-2}}, \frac{[D]^{1/2}+P_{k-1}}{Q_{k-1}},$$

the first k complete quotients of the continued fraction (3.12) *are*

$$\frac{[D]^{1/2}+P_k}{Q_{k-1}}, \frac{[D]^{1/2}+P_{k-1}}{Q_{k-2}}, \ldots, \frac{[D]^{1/2}+P_2}{Q_1}, \frac{[D]^{1/2}+P_1}{Q_0}.$$

The proof of Theorem 3.2 bears a close resemblance to that of an analogous theorem [3, p. 84], and is left to the reader.

The elements of a pure periodic proper continued fraction must all be positive, and, accordingly, all the complete quotients y_n must be positive. This statement cannot be made about the conjugate numbers z_n, but these have a number of interesting properties some of which are described in the Theorems 3.3-3.9.

THEOREM 3.3. *If a conjugate number z_{n_0} is negative, all of the conjugate numbers z_n $(n > n_0)$ are negative.*

The proof of this theorem is strictly analogous to that for the corresponding theorem for regular continued fractions [3, p. 81].

THEOREM 3.4. *The conjugate of the value of a pure periodic proper continued fraction is a negative number.*

Theorem 3.4 is an immediate consequence of Theorem 3.1.

THEOREM 3.5. *There exists a positive integer n_0 such that $z_n < 0$ $(n > n_0)$, where z_n is the conjugate of the n-th complete quotient of the periodic proper continued fraction*

$$(3.51) \quad b_0 + \frac{a_1}{b_1} + \cdots + \frac{a_{h-1}}{b_{h-1}} + \frac{a_h}{b_h} + \frac{a_{h+1}}{b_{h+1}} + \cdots + \frac{a_{h+k-1}}{b_{h+k-1}} + \frac{a_{h+k}}{b_h} + \frac{a_{h+1}}{b_{h+1}} + \cdots.$$

If the continued fraction (3.51) has a preliminary period of h terms, it follows from Theorem 3.4 that $n_0 = h - 1$ is a number which satisfies the conditions of Theorem 3.5.

The following theorems give a more precise criterion for the determination of the signs of the conjugate numbers.

THEOREM 3.6. *If z_0, the conjugate of the value y_0 of the proper continued fraction* (3.51), *satisfies the condition*

(3. 61) $$b_0 < z_0 < y_0,$$

there then exists a unique positive integer m such that the conditions

(3. 62) $$\frac{A_{2m-2}}{B_{2m-2}} < z_0 < \frac{A_{2m}}{B_{2m}},$$

(3. 63) $$z_{2m+1} < 0 < z_{2m},$$

both hold.

The convergents with even subscripts are less than y_0, and each one is a better approximation to y_0 than those which precede it [2]. Thus, since the convergents are rational numbers, and z_0 is irrational, there exists a unique integer m such that condition (3. 62) holds.

The equality

(3. 64) $$y_0 = \frac{y_nA_{n-1} + a_nA_{n-2}}{y_nB_{n-1} + a_nB_{n-2}} \qquad (n = 2, 3, 4, \cdots)$$

will still be valid if z_n is substituted for y_n, and z_0 for y_0. The formula, which is obtained by this substitution, may be written in the form

(3. 65) $$z_n = -a_n \frac{B_{n-2}}{B_{n-1}} \cdot \frac{z_0 - \dfrac{A_{n-2}}{B_{n-2}}}{z_0 - \dfrac{A_{n-1}}{B_{n-1}}} \qquad (n = 2, 3, 4, \cdots).$$

Since each odd convergent is great than y_0 it follows from condition (3. 61) that

$$z_0 - \frac{A_{2m-1}}{B_{2m-1}} < 0 \qquad (m = 1, 2, 3, \cdots).$$

Hence, condition (3. 63) is an immediate consequence of conditions (3. 62) and (3. 65).

THEOREM 3. 7. *The conditions*

(3. 71) $$z_0 < b_0,$$

(3. 71) $$z_1 < 0,$$

are equivalent.

The equivalence of conditions (3. 71) and (3. 72) is apparent when it is recalled [2] that

(3. 73) $$z_1 = \frac{a_1}{z_0 - b_0}.$$

THEOREM 3.8. *If z_0, the conjugate of the value y_0 of the continued fraction (3.51), satisfies the condition*

$$y_0 < z_0 < b_0 + \frac{a_1}{b_1}$$

there then exists a unique positive integer m such that the conditions

$$\frac{A_{2m+1}}{B_{2m+1}} < z_0 < \frac{A_{2m-1}}{B_{2m-1}},$$
$$z_{2m+2} < 0 < z_{2m+1},$$

both hold.

The proof of Theorem 3.8 is strictly analogous to that of Theorem 3.6.

THEOREM 3.9. *The conditions*

$$b_0 + \frac{a_1}{b_1} < z_0, \tag{3.91}$$

$$z_2 < 0, \tag{3.92}$$

are equivalent.

Condition (3.91) may be expressed in the form

$$a_1 < b_1 z_0 - b_0 b_1, \tag{3.93}$$

and the equation (3.73) may be written as

$$z_1 = \frac{a_1 b_1}{b_1 z_0 - b_0 b_1}. \tag{3.94}$$

It follows from (3.93) and (3.94) that $z_1 < b_1$, and since

$$z_2 = \frac{a_2}{z_1 - b_1},$$

it is seen that (3.92) holds. The steps in the above proof are reversible, and thus it is clear that the conditions (3.91) and (3.92) are equivalent.

4. Reduced quadratic irrationals. A quadratic irrational is said to be a *reduced* quadratic irrational if it is greater than 1 and its conjugate lies between -1 and 0.

It is well-known that every pure periodic regular continued fraction represents a reduced quadratic irrational [3, p. 80]. This statement cannot be made, without qualification, about pure periodic proper continued fractions, as is shown in the next theorem.

THEOREM 4. 1. *The pure periodic proper continued fraction*

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_k}{b_0} + \frac{a_1}{b_1} + \cdots \tag{4.11}$$

represents a reduced quadratic irrational y_0 *if* $b_{k-1} \geqq a_k$. *It does not represent a reduced quadratic irrational if* $b_{k-1} < a_k$ *and* $b_{k-2} \geqq a_{k-1}$.

It will be recalled that it was shown in Theorem 2. 1 that the continued fraction (2. 12) converges to the value $-z_0$. The n-th complete quotient of the continued fraction (2. 12) will be represented by w_n. If $b_{k-1} \geqq a_k$ it is clear that $w_1 > a_k$, and hence $w_0 < 1$. This implies that $-1 < z_0 < 0$, and, since the continued fraction (4. 11) is a proper continued fraction, $1 \leqq a_k \leqq b_0 < y_0$. Thus, if $b_{k-1} \geqq a_k$, the value y_0 of (4. 11) is a reduced quadratic irrational. On the other hand, if $b_{k-1} < a_k$, and if $b_{k-2} \geqq a_{k-1}$, then

$$w_2 > a_{k-1}, \qquad 1 > a_{k-1}/w_2, \qquad w_1 < a_k, \qquad \text{and} \qquad w_0 = a_k/w_1 > 1.$$

Thus $z_0 < -1$, and y_0 is not a reduced quadratic irrational

An example of a pure periodic proper continued fraction whose value is not a reduced quadratic irrational is the continued fraction

$$6 + \frac{1}{2} + \frac{4}{6} + \frac{1}{2} + \frac{4}{6} + \cdots$$

whose value is $([273]^{1/2} + 9)/4$.

All the complete quotients of a pure periodic regular continued fraction are reduced quadratic irrationals. Such need not be the case for pure periodic proper continued fractions, but there is a sub-class of these continued fractions whose members do have this property.

THEOREM 4. 2. *If the elements of the pure periodic proper continued fraction* (4. 11) *satisfy the conditions*

$$b_n \geqq a_{n+1} \qquad (n = 0, 1, 2, \cdots), \tag{4.21}$$

all the complete quotients of (4. 11) *are reduced quadratic irrationals.*

Theorem 4. 2 is an immediate consequence of Theorem 4. 1. It should be noted that pure periodic proper continued fractions, whose partial numerators are all equal to the same positive integer, satisfy conditions (4. 21) since it is characteristic of proper continued fractions that $b_n \geqq a_n$ $(n = 1, 2, 3, \cdots)$.

No proofs will be given of the theorems which follow in this section since they follow closely the proofs for the corresponding theorems for regular continued fractions [3, pp. 81-92].

THEOREM 4. 3. *If y_0 is the value of the periodic proper continued fraction*

$$(4.31) \quad b_0 + \frac{a_1}{b_1} + \cdots + \frac{a_{k-1}}{b_{k-1}} + \frac{a_k}{b_k} + \frac{a_{k+1}}{b_{k+1}} + \cdots + \frac{a_{k+k-1}}{b_{k+k-1}} + \frac{a_{k+k}}{b_k} + \frac{a_{k+1}}{b_{k+1}} + \cdots$$

where

$$(4.32) \qquad a_{n+k} = a_n \qquad (n = 1, 2, 3, \cdots),$$

and the elements satisfy condition (4. 21), *and if the conjugate z_0 of y_0 is a negative number, the continued fraction* (4. 31) *has at most one non-periodic term.*

THEOREM 4. 4. *If, in addition to the conditions of Theorem* 4. 3, *it is required that z_0 be greater than* -1, *then the proper continued fraction* (4. 31) *will be pure periodic.*

THEOREM 4. 5. *If the elements of the periodic proper continued fraction* (4. 31) *satisfy the conditions* (4. 21) *and* (4. 32), *then*

(a) *if* $-1 < z_0 < 0$, *the continued fraction* (4. 31) *is pure periodic,*

(b) *if* $z_0 < -1$, (4. 31) *has one non-periodic term,*

(c) *if* $0 < z_0$, (4. 31) *has one or more non-periodic terms.*

THEOREM 4. 6. *The value of the periodic continued fraction*

$$(4.61) \quad b_0 + \frac{a}{b_1} + \frac{a}{b_2} + \cdots + \frac{a}{b_2} + \frac{a}{b_1} + \frac{a}{2b_0} + \frac{a}{b_1} + \frac{a}{b_2} + \cdots,$$

where all the partial numerators are equal to the same positive integer and

$$b_n = b_{n+k} \qquad (n = 1, 2, 3, \cdots),$$

$$b_n = b_{k-n} \qquad (n = 1, 2, 3, \cdots, k-1),$$

is the square root of a rational number.

For example, the value of the continued fraction

$$1 + \frac{2}{3} + \frac{2}{4} + \frac{2}{3} + \frac{2}{2} + \frac{2}{3} + \frac{2}{4} + \frac{2}{2} + \frac{2}{3} + \cdots$$

is $[5/2]^{1/2}$.

THEOREM 4. 7. *If the complete quotients of the continued fraction* (4. 61) *are written in the form*

$$\frac{[D]^{1/2}}{Q_0}, \frac{[D]^{1/2} + P_1}{Q_1}, \frac{[D]^{1/2} + P_2}{Q_2}, \cdots, \frac{[D]^{1/2} + P_k}{Q_k},$$

where D is a positive integer, not a perfect square, then

$$P_{n+1} = P_{k-n} \qquad (n = 0, 1, 2, \cdots, k-1)$$

$$Q_n = Q_{k-n} \qquad (n = 0, 1, 2, \cdots, k).$$

THEOREM 4. 8. *If the primitive period of the continued fraction* (4. 61) *contains k terms, then $P_n = P_{n+1}$ only when k is an even integer, and $n = k/2$, while $Q_n = Q_{n+1}$ only when k is odd, and $n = (k-1)/2$.*

Theorems 4. 6-4. 8 may be illustrated by the continued fraction

$$8 + \frac{3}{5+} \, \frac{3}{5+} \, \frac{3}{3+} \, \frac{3}{4+} \, \frac{3}{3+} \, \frac{3}{5+} \, \frac{3}{5+} \, \frac{3}{16+} \, \frac{3}{5+} \cdots.$$

Here, $k = 8$, and the complete quotients are

$$y_0 = \frac{[73]^{1/2}}{1}, \qquad y_5 = \frac{[73]^{1/2} + 6}{37/9},$$

$$y_1 = \frac{[73]^{1/2} + 8}{3}, \qquad y_6 = \frac{[73]^{1/2} + 19/3}{8/3},$$

$$y_2 = \frac{[73]^{1/2} + 7}{8/3}, \qquad y_7 = \frac{[73]^{1/2} + 7}{3},$$

$$y_3 = \frac{[73]^{1/2} + 19/3}{37/9}, \qquad y_8 = \frac{[73]^{1/2} + 8}{1}.$$

$$y_4 = \frac{[73]^{1/2} + 6}{3},$$

5. Periodic proper continued fraction expansions.

In the preceding sections the emphasis was placed upon the continued fraction itself, and not upon its value. In this section the continued fraction will be considered as a representation or expansion of the number which is its value.

The properties of proper continued fraction expansions in the case where the members of the sequence $\{a_n\}$ are all equal to unity, have been studied in considerable detail. For example, it is well-known that not only does every periodic regular continued fraction represent a quadratic irrational, but the regular continued fraction expansion of a quadratic irrational is periodic. Such a result would not be expected to hold in general for proper continued fraction representations of quadratic irrationals. This is clear when it is recalled that if there is a monotone increasing subsequence among the partial numerators, then there will also exist such a subsequence among the partial denominators, and hence periodicity will not occur [Cf. 2]. Thus it is apparent that the question of the periodicity of a given expansion is dependent not only upon the number which is represented by the expansion, but also upon the sequence of partial numerators which is used.

The question arises as to what quadratic irrationals admit periodic proper continued fraction expansions other than a regular continued fraction expansion. This question is answered, in part, in the theorem which follows.

THEOREM 5. 1. *Every quadratic irrational, which admits a pure periodic regular continued fraction expansion with an odd number of terms in the period, may be expressed in an infinitude of ways as a proper periodic continued fraction, the partial numerators being equal to the same positive integer, and the period consisting of two terms which are preceded by two terms.*

Represent by y_0 the quadratic irrational, and write its regular continued fraction expansion in the form

$$d_0 + \frac{1}{d_1} + \frac{1}{d_2} + \cdots + \frac{1}{d_{h-1}} + \frac{1}{d_0} + \frac{1}{d_1} + \cdots, \tag{5. 11}$$

where h, an odd positive integer, is the number of terms in the period. Let C_n/D_n be the n-th convergent of the expansion.

The continued fraction whose n-th convergent is A_n/B_n, where

$$A_n = C_{nk}, \qquad B_n = D_{nk} \qquad (n = 0, 1, 2, \cdots),$$

and k is an odd multiple of h, may now be constructed [3, p. 200]. This continued fraction is equivalent to the continued fraction

$$d_0 + \frac{D_{k-1,1}}{D_k} + \frac{D_{k-1,1}}{C_{k-1,1} + D_{k-2,1}} + \frac{D_{k-1,1}}{D_{2k-1,1}} + \cdots. \tag{5. 12}$$

By substituting for k in (5. 51) successive odd multiples of h the sequence of continued fractions referred to in the statement of Theorem 5. 5 may be obtained.

It remains to show that the continued fractions obtained in this way are proper continued fractions. This is done by referring to the relations

$$\begin{aligned} D_k &= d_1 D_{k-1,1} + D_{k-2,2}, \\ D_{2k-1,1} &= D_{k-1,1}\,(C_{k-1,1} + D_{k-2,1}), \\ C_{k-1,1} + D_{k-2,1} &= D_k + D_{k-2,1}, \end{aligned}$$

which were obtained from the fundamental formulas [3, p. 14]. The proof of Theorem 5. 1 is complete.

It is interesting to note that the proof of Theorem 5. 1 is independent of the length of the regular continued fraction so long as this period is odd. The theorem may be illustrated by considering the continued fraction

$$1+\frac{1}{1+}\,\frac{1}{1+}\cdots$$

whose value is $([5]^{1/2}+1)/2$. In this case $h=1$, and k may be any odd positive integer. The number $([5]^{1/2}+1)/2$ admits an expansion of the kind described in Theorem 5.1 if the members of the sequence $a_1, a_2, a_3, \cdots$ are all equal to any one of the numbers 2, 3, 5, 13, 34, 89, 233, 610, 1597, 4181, 10946, 28657, 75025, 196418, $\cdots$. (There are integers other than those in the indicated sequence which result in periodic expansions). Three expansions are

$$1+\frac{2}{3+}\,\frac{2}{4+}\,\frac{2}{8+}\,\frac{2}{4+}\,\frac{2}{8+}\cdots,$$

$$1+\frac{5}{8+}\,\frac{5}{55+}\,\frac{5}{11+}\,\frac{5}{55+}\,\frac{5}{11+}\cdots,$$

$$1+\frac{233}{377+}\,\frac{233}{121{,}393+}\,\frac{233}{521+}\,\frac{233}{121{,}393+}\,\frac{233}{521+}\cdots.$$

It is clear that Theorem 5.1 is also valid for quadratic irrationals of the form w_0+I where I is an integer, and w_0 is a reduced quadratic irrational with an odd number of terms in its period.

6. Conditions that a fraction be equal to a convergent. Before sufficient conditions are obtained that a fraction be equal to a convergent of a given proper continued fraction, the following necessary condition will be established.

THEOREM 6.1. *If*

$$b_0+\frac{a_1}{b_1+}\,\frac{a_2}{b_2+}\cdots$$

is a proper continued fraction expansion of the real number y_0,

$$\left|y_0-\frac{A_n}{B_n}\right|<\frac{a_1a_2\cdots a_n}{B_n(B_{n-1}+B_n)} \qquad (n=1,2,3,\cdots). \tag{6.11}$$

In the proof of Theorem 6.1 use will be made of the well-known formulas

$$y_0=\frac{y_{n+1}A_n+a_{n+1}A_{n-1}}{y_{n+1}B_n+a_{n+1}B_{n-1}}, \tag{3.64}$$

$$A_nB_{n-1}-A_{n-1}B_n=(-1)^{n-1}a_1a_2\cdots a_n \qquad (n=1,2,3,\cdots).$$

It is clear that

$$y_0-\frac{A_n}{B_n}=-\frac{a_{n+1}(A_nB_{n-1}-A_{n-1}B_n)}{B_n(y_{n+1}B_n+a_{n+1}B_{n-1})}=\frac{(-1)^n a_1a_2\cdots a_{n+1}}{B_n(y_{n+1}B_n+a_{n+1}B_{n-1})}$$
$$(n=1,2,3,\cdots).$$

Since, by (5.1), $y_{n+1}\geqq b_{n+1}$, it follows that

$$\left| y_0 - \frac{A_n}{B_n} \right| \leqq \frac{a_1 a_2 \cdots a_{n+1}}{B_n B_{n+1}}.$$

By means of the recursion formulas (2.13), it may be verified that

$$\frac{a_{n+1}}{B_{n+1}} \leqq \frac{1}{B_{n-1} + B_n} \qquad (n = 1, 2, 3, \cdots),$$

and, hence,

$$(6.12) \qquad \left| y_0 - \frac{A_n}{B_n} \right| \leqq \frac{a_1 a_2 \cdots a_{n+1}}{B_n B_{n+1}} \leqq \frac{a_1 a_2 \cdots a_n}{B_n (B_{n-1} + B_n)} \quad (n = 1, 2, 3, \cdots).$$

The first equality sign of (6.12) applies only if $y_{n+1} = b_{n+1}$, but, in this eventuality, the second equality sign cannot hold since b_{n+1} is then greater than a_{n+1}. Thus, Theorem 6.1 is established.

It will be useful in what follows to define what is meant by a *modified* terminating proper continued fraction. Such a continued fraction is obtained by replacing the final partial numerator b_n by $(b_n - 1) + a_{n+1}/b_{n+1}$ where $a_{n+1} = b_{n+1}$. It will be noted that this substitution does not change the value of the continued fraction.

THEOREM 6.2. *Let*

$$b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots$$

be a proper continued fraction expansion of the number $y_0 \neq b_0$, *and let*

$$6.21) \qquad c_0 + \frac{a_1}{c_1 +} \frac{a_2}{c_2 +} \cdots \frac{a_n}{+ c_n} \qquad (n \geqq 1)$$

be the (possibly modified) proper continued fractional expansion of a rational number $p/q \neq y_0$, *where q is a positive integer, and where n is chosen in such a fashion that* $(-1)^n$ *has the same sign as* $(y_0 - p/q)$. *A necessary and sufficient condition that p/q be equal to the n-th convergent of the expansion of* y_0 *is that*

$$(6.22) \qquad \left| y_0 - \frac{p}{q} \right| < \frac{a_1 a_2 \cdots a_n}{Y_n (Y_{n-1} + Y_n)}$$

where Y_n *is the denominator of the n-th convergent of the continued fraction* (6.21).

The necessity of condition (6.22) was established in Theorem 6.1. The sufficiency of this condition will now be established. Let X_n represent the numerator of the n-th convergent of the continued fraction (6.21). Since $Y_n y_0 - X_n \neq 0$, the equation

$$(6.23) \qquad y_0 = \frac{w X_n + a_{n+1} X_{n-1}}{w Y_n + a_{n+1} Y_{n-1}}$$

defines a real number w. If X_n/Y_n is subtracted from equation (6. 23) the equalities

$$y_0 - \frac{X_n}{Y_n} = -\frac{a_{n+1}(X_nY_{n-1} - X_{n-1}Y_n)}{Y_n(wY_n + a_{n+1}Y_{n-1})} = \frac{(-1)^n a_1 a_2 \cdots a_{n+1}}{Y_n(wY_n + a_{n+1}Y_{n-1})}$$

are obtained. Hence

$$\left| y_0 - \frac{X_n}{Y_n} \right| = \frac{a_1 a_2 \cdots a_{n+1}}{Y_n(wY_n + a_{n+1}Y_{n-1})}, \tag{6.24}$$

since $(-1)^n$ and $(y_0 - X_n/Y_n)$ have the same sign. If equation (6. 24) is combined with the inequality (6. 22), the inequality

$$\frac{a_{n+1}}{wY_n + a_{n+1}Y_{n-1}} < \frac{a_{n+1}}{a_{n+1}Y_n + a_{n+1}Y_{n-1}} \tag{6.25}$$

results. The inequality (6. 25) indicates that w is greater than a_{n+1}. Equation (6. 23) may be written in the form

$$y_0 = c_0 + \frac{a_1}{c_1 +}\,\frac{a_2}{c_2 +} \cdots \frac{a_n}{+ c_n +}\,\frac{a_{n+1}}{w}.$$

Denote by

$$c_{n+1} + \frac{a_{n+2}}{c_{n+2} +}\,\frac{a_{n+3}}{c_{n+3} +} \cdots$$

the proper continued fraction expansion of w corresponding to the integers $a_{n+2}, a_{n+3}, a_{n+4}, \cdots$. Since w is greater than a_{n+1}, $c_{n+1} \geqq a_{n+1}$, and

$$c_0 + \frac{a_1}{c_1 +}\,\frac{a_2}{c_2 +} \cdots \frac{a_n}{+ c_n +}\,\frac{a_{n+1}}{c_{n+1} +}\,\frac{a_{n+2}}{c_{n+2} +} \cdots$$

is a proper continued fraction expansion of y_0. But the proper continued fraction expansion of a number corresponding to a given set of positive integers is unique [2]. Hence, $c_n = b_n$ for all non-negative integral values of n, and p/q is equal to the n-th convergent of the proper continued fraction expansion of y_0 corresponding to the sequence of positive integers $a_1, a_2, a_3, \cdots$. This completes the proof of the sufficiency of condition (6. 22).

A corollary may now be stated.

COROLLARY 6. 21. *If a rational number p/q satisfies the conditions of Theorem* 6. 2 *save perhaps condition* (6. 22), *and if*

$$\left| y_0 - \frac{p}{q} \right| < \frac{a_1 a_2 \cdots a_n}{2Y_n^2}$$

then p/q is equal to the n-th convergent of the proper continued fraction expansion of the number y_0, corresponding to the sequence of positive integers $a_1, a_2, a_3, \cdots$.

The corollary will be established by showing that condition (6.22) is satisfied. It may be shown, with the help of the recursion formulas (2.13), that

$$(6.27)\qquad Y_0 \leqq Y_1 < Y_2 < \cdots < Y_n < \cdots \qquad (n = 1, 2, 3, \cdots),$$

and hence that

$$Y_n(Y_n + Y_{n-1}) \leqq 2Y_n^2.$$

Thus

$$\left| y_0 - \frac{p}{q} \right| < \frac{a_1a_2\cdots a_n}{2Y_n^2} \leqq \frac{a_1a_2\cdots a_n}{Y_n(Y_n + Y_{n-1})},$$

and (6.22) is satisfied.

The final theorem of this section contains a third sufficient condition.

THEOREM 6.3. *If y_0 is a positive number, not an integer; if*

$$(6.31)\qquad | p^2 - y_0^2q^2 | < a_1a_2\cdots a_ny_0 \qquad (n > 0),$$

$$(6.32)\qquad \frac{p}{q} = b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_n}{b_n} \qquad (n, p, q > 0),$$

n being selected, by modifying the expansion if necessary, so that $(-1)^n$ has the same sign as $(y_0 - p/q)$, and if

$$(6.33)\qquad p = X_n, \qquad q = Y_n \qquad (n > 0)$$

where X_n/Y_n is the last convergent of the proper continued fraction (6.32), then p/q is the n-th convergent of the proper continued fraction expansion of y_0 corresponding to the sequence of positive integers

$$\{c_i\} \qquad (i = 1, 2, 3, \cdots),$$

where

$$c_i = a_i \qquad (i = 1, 2, 3, \cdots, n),$$

and the remaining elements of the sequence are arbitrary.

Consider the inequality

$$(6.34)\qquad \frac{(-1)^n\theta a_1a_2\cdots a_n}{y_0Y_n + X_n} < 1, \qquad 0 < \theta < 1.$$

The inequality (6.34) is valid when $n > 1$, since it may be shown by a simple induction that

$$a_1a_2\cdots a_n < X_n \qquad (n > 1).$$

The inequalities (6.27) may be written in the form

$$Y_{n-1} + 1 \leqq Y_n \qquad (n > 1),$$

and, when this result is combined with the inequality (6.34) it follows that

$$(6.35) \qquad y_0Y_{n-1} < y_0Y_n - \frac{(-1)^n a_1a_2\cdots a_n\theta y_0}{y_0Y_n + X_n}.$$

It is obvious that the inequality (6.35) holds when $n = 1$. By hypothesis

$$(6.36) \qquad y_0Y_n - X_n = \frac{(-1)^n\theta a_1a_2\cdots a_n y_0}{y_0Y_n + X_n} \qquad (n > 0).$$

If the inequality (6.35) is added to the equation (6.36), it is found that

$$y_0(Y_{n-1} + Y_n) < y_0Y_n + X_n \qquad (n > 0).$$

The weaker inequality

$$\theta y_0(Y_{n-1} + Y_n) < y_0Y_n + X_n$$

is valid when $n = 0$, and is equivalent to the inequality

$$(6.37) \qquad \frac{a_1a_2\cdots a_n\theta y_0}{Y_n(y_0Y_n + X_n)} < \frac{a_1a_2\cdots a_n}{Y_n(Y_{n-1} + Y_n)}.$$

Equation (6.36) may be written in the form

$$(6.38) \qquad \frac{a_1a_2\cdots a_n\theta y_0}{Y_n(y_0Y_n + X_n)} = \left| y_0 - \frac{X_n}{Y_n} \right|.$$

Conditions (6.37) and (6.38) are equivalent to condition (6.22), and, thus, p/q is the n-th convergent of y_0.

THE RICE INSTITUTE.

REFERENCES

1. D. H. Lehmer, "A cotangent analog of continued fractions," *Duke Mathematical Journal*, vol. 4 (1938), pp. 323-340.
2. W. Leighton, "Proper continued fractions," *American Mathematical Monthly*, vol. 47 (1940), pp. 274-280.
3. O. Perron, *Die Lehre von den Kettenbrüchen*, zweite Auflage (Leipzig), 1929.

ON THE Ω-MATRICES OF TOEPLITZ.*

By Aurel Wintner.

The theory of Ω-matrices was developed by Toeplitz[1] for the case of analytic Fourier series (Laurent series). The restriction of analyticity is essential for the method of Toeplitz, since use is made of Cauchy's theorem. The object of the present note is to develop the theory in its full generality, that is, for the case of those measurable functions for which such a theory is possible at all.

Due to such an approach, some of the results turn out to hold even in the almost periodic case of relative measure.

1. By a distribution function (mod 1) will be meant a monotone non-decreasing function $\phi(e^{2\pi ix})$, $0 \leqq x \leqq 1$, of total variation 1. Two such functions, say ϕ_1 and ϕ_2, will be considered as identical if and only if

$$(1) \qquad \theta(x+0) = \text{const.} = \theta(x-0),$$

where $\theta(x) = \phi_1(e^{2\pi ix}) - \phi_2(e^{2\pi ix})$. In this sense, the trigonometric momenta

$$(2) \qquad \lambda_n(\phi) \equiv \int_0^1 e^{2\pi nix} d\phi(e^{2\pi ix}) = \bar{\lambda}_{-n}(\phi)\,;\ n = 1, 2, \cdots, (\lambda_0 = 1),$$

determine ϕ uniquely, since, as a consequence of the trigonometric approximation of Weierstrass, every function $\theta(x)$, $0 \leqq x \leqq 1$, of bounded variation must satisfy (1) if it satisfies

$$(3) \qquad \int_0^1 e^{2\pi nix} d\theta(x) = 0 \text{ for } n = 0, \pm 1, \pm 2, \cdots.$$

According to the Herglotz formulation of the Toeplitz criterion of real harmonic functions which are regular and non-negative within the unit circle, there exists, for given numbers λ_n, a distribution function ϕ satisfying (2) if and only if

$$(4) \qquad \sum_{n=0}^{k}\sum_{m=0}^{k} \lambda_{n-m}\xi_n\bar{\xi}_m \geqq 0, \qquad (\lambda_0 = 1),$$

where k and the ξ are arbitrary.

---

* Received April 19, 1941.

[1] O. Toeplitz, "Zur Theorie der quadratischen und bilinearen Formen von unendlich vielen Veränderlichen, I: Theorie der *L*-Formen," *Mathematische Annalen*, vol. 70 (1911), pp. 351-376, more particularly § 7-§ 8.

Consider the class of those measurable functions $f = f(e^{2\pi it})$, $0 \leqq t \leqq 1$, which satisfy

$$(5) \qquad |f(e^{2\pi it})| = 1, \text{ i. e. } \bar{f}(e^{2\pi it}) = 1/f(e^{2\pi it}), \text{ for every } t.$$

The n-th power, $f^n = f(e^{2\pi it})^n$, of such an f belongs to the same class for $n = 0, \pm 1, \pm 2, \cdots$, and has the Fourier series

$$(6) \qquad f(e^{2\pi it})^n \sim \sum_{m=-\infty}^{\infty} \omega_{nm}(f) e^{2\pi imt},$$

where

$$(7) \qquad \omega_{nm}(f) = \int_0^1 f(e^{2\pi it})^n e^{-2\pi imt}\, dt, \qquad (n, m = 0, \pm 1, \pm 2, \cdots).$$

In particular

$$(8) \qquad \omega_{n0}(f) = \int_0^1 f(e^{2\pi it})^n\, dt = \bar{\omega}_{-n0}(f); \qquad \omega_{00}(f) = 1,$$

and so (4) is satisfied by $\lambda_n = \omega_{n0}(f)$; in fact, from (5),

$$\sum_{n=0}^{k} \sum_{m=0}^{k} \int_0^1 f(e^{2\pi it})^{n-m}\, dt\, \xi_n \bar{\xi}_m = \int_0^1 \Big| \sum_{n=0}^{k} f(e^{2\pi it})^n \xi_n \Big|^2 dt \geqq 0,$$

where k and the ξ are arbitrary. Thus there exists a distribution function (mod 1) by means of which (8) is representable in the form (2); so that

$$(9) \qquad \int_0^1 f(e^{2\pi it})^n\, dt = \int_0^1 e^{2\pi inx}\, d\phi(e^{2\pi ix}); \qquad n = 0, \pm 1, \pm 2, \cdots.$$

On writing the Lebesgue integral on the left of (9) as a Stieltjes integral, one sees, from the uniqueness theorem of the trigonometric momentum problem, that $\phi(e^{2\pi ix})$ represents the Lebesgue measure of that subset of the interval $0 \leqq t < 1$ on which $\arg f(e^{2\pi it}) < 2\pi x$, where $\arg f(t)$ and x are thought of as reduced mod 1. Correspondingly, $\phi(e^{2\pi ix})$ will be called the distribution function of $f(e^{2\pi it})$.

2. In the sequel, $A = \| a_{nm} \|$ will denote an infinite matrix in which both n and m run through the values $0, \pm 1, \pm 2, \cdots$. The matrix will be said to be bounded if the corresponding ∞-ary bilinear form is bounded in the sense of Hilbert. The matrix A will be called non-singular if it is bounded and such that there exists a bounded matrix, B, satisfying

$$AB = E \text{ and } BA = E,$$

where E is the unit matrix. Then, according to Toeplitz, B is uniquely determined by A, and can therefore be denoted by A^{-1}.

If ϕ is a distribution function (mod 1), let $\Lambda(\phi)$ denote the matrix

$$\Lambda(\phi) = \| \lambda_{n-m}(\phi) \|, \tag{10}$$

where the λ are the trigonometric momenta (2).

If $f(e^{2\pi it})$, $0 \leqq t \leqq 1$, is a measurable function satisfying (5), let $\Omega(f)$ denote the matrix

$$\Omega(f) = \| \omega_{nm}(f) \|, \tag{11}$$

where the ω are the Fourier constants (7).

For these two types of matrices, the following facts will first be established:

(I) *The matrix* $\Lambda(\phi)$ *is bounded if and only if* $\phi = \phi(e^{2\pi ix})$, $0 \leqq x \leqq 1$, *satisfies a uniform Lipschitz condition.*

It is understood that a function $F(s)$ is said to satisfy a uniform Lipschitz condition if

$$| F(s_1) - F(s_2) | \leqq \text{Const.} \, | s_1 - s_2 |, \tag{12}$$

where s_1, s_2 are two arbitrary points of the s-range on which F is defined. Since

$$0 < \text{const.} < \frac{| e^{i\beta} - e^{i\alpha} |}{\beta - \alpha} < \text{Const. for } 0 \leqq \alpha < \beta < 2\pi, \tag{13}$$

it is immaterial whether the uniform Lipschitz condition of (I) is meant to refer to $s = e^{2\pi ix}$ or to $s = x$.

(II) *The matrix* $\Omega(f)$, *which need not be bounded, is always such that* $\Omega(f)\Omega^*(f)$ *exists and is identical with* $\Lambda(\phi)$, *where* ϕ *denotes the distribution function of* f.

It is understood that $\Omega(f)\Omega^*(f)$ denotes the matrix product AA^*, where $A = \Omega(f)$, and that A^* is defined as the transposed matrix of the complex conjugate, $\bar{A}$, of the matrix A.

(III) *The matrix* $\Omega(f)$ *is bounded if and only if the distribution function,* $\phi = \phi(e^{2\pi ix})$, *of* $f = f(e^{2\pi it})$ *satisfies a uniform Lipschitz condition.*

A function $F = F(s)$ will be said to be of bounded distortion if

$$\text{const.} \, | s_2 - s_1 | \leqq | F(s_2) - F(s_1) | \leqq \text{Const.} \, | s_2 - s_1 |, \text{ where const.} > 0, \tag{14}$$

holds for any pair s_1, s_2 of points of the s-range on which F is defined. If F is a function of $e^{2\pi it}$ alone, (13) shows that (14) is the same condition for $s = e^{2\pi it}$ as for $s = t$.

(IV) *The matrix* $\Lambda(\phi)$ *is non-singular if and only if the distribution function,* $\phi = \phi(e^{2\pi i x})$, *of* $f = f(e^{2\pi i t})$ *is of bounded distortion.*

(IV) becomes false if one replaces $\Lambda(\phi)$ by $\Omega(f)$. This is shown by the example $f(e^{2\pi i t}) = e^{2\pi l i t}$, which has indeed the distribution function $\phi = \text{const.}$ $(= 1)$ for every l $(= 1, 2, \cdots)$ but determines a bounded matrix $\Omega(f)$ only when $l = 1$. All that can be said in this direction may be formulated as follows:

(V) *In order that* $\Omega(f)$ *be non-singular, it is necessary but not sufficient that* $\Omega^*(f)\Omega(f)$ *be non-singular.*

In the proof of these statements, use will be made, besides the matrix (10) defined by (2), of the ordinary Laurent matrix,

$$L(g) = \| \int_0^1 g(e^{2\pi i t}) e^{2\pi i (n-m) t}\, dt \|, \tag{15}$$

where $g = g({}^{2\pi i t})$ is any Lebesgue integrable function. If g_1, g_2 and their squares are integrable in the sense of Lebesgue, then, by the polarized form of Parseval's relation,

$$L(g_1 g_2) = L(g_1)L(g_2) = L(g_2 g_1). \tag{16}$$

It is clear also from (15) that

$$L(1) = E, \qquad (E = \text{unit matrix}), \tag{17}$$

and

$$L(g) = L^*(g). \tag{18}$$

According to Toeplitz, $L(g)$ is bounded if and only if

$$| g(e^{2\pi i t}) | < \text{Const. almost everywhere}, \tag{19}$$

and $L(g)$ is non-singular if and only if

$$0 < \text{const.} < | g(e^{2\pi i t}) | < \text{Const. almost everywhere.} \tag{20}$$

3. In order to prove (I), we first observe that, if a matrix $A = \| a_{nm} \|$ is bounded, then

$$\sum_{m=-\infty}^{\infty} | a_{nm} |^2 < \infty \qquad \text{for } n = 0, \pm 1, \pm 2, \cdots.$$

Hence

$$\sum_{n=-\infty}^{\infty} | \lambda_n(\phi) |^2 < \infty$$

is a necessary condition for the boundedness of the matrix (11). It follows, therefore, from the Fischer-Riesz theorem that there exists a function $\delta = \delta(x)$, $0 \leqq x \leqq 1$, of class (L) such that

$$\lambda_n(\phi) = \int_0^1 \delta(x) e^{2\pi i n x}\, dx \qquad \text{for } n = 0, \pm 1, \pm 2, \cdots$$

(the fact that δ actually is of class (L^2) will not be needed). On comparing this with (2), one sees that (3) is satisfied by

$$\theta(x) = \int_0^x \delta(y)\, dy - \phi(e^{2\pi i x}).$$

Since this $\theta(x)$ is of bounded variation, so that (3) implies (1), it follows from the fundamental theorem of calculus that the derivative, $\phi'(e^{2\pi i x})$, of the monotone function ϕ of $e^{2\pi i x}$ is identical with $\delta(x)$ almost everywhere. Thus $\phi(e^{2\pi i x})$ is absolutely continuous and

$$(21) \qquad \Lambda(\phi) = L(\phi'),$$

by (10), (2) and (15). Since (19) is necessary and sufficient for the boundedness of (15), it follows that $\Lambda(\phi)$ is bounded if and only if $|\,\phi'\,| <$ Const. almost everywhere. Hence, in order to complete the proof of (I), it is sufficient to realize that a function satisfies a uniform Lipschitz condition if and only if it is absolutely continuous and has a derivative (almost everywhere) which is bounded on the complement of a zero set.

Next, if $f(e^{2\pi i t})$ is a measurable function satisfying (5), then, by the polarized form of Parseval's relation

$$\sum_{k=-\infty}^{\infty} \int_0^1 f(e^{2\pi i t})^n e^{-2\pi i k t}\, dt. \int_0^1 \bar{f}(e^{2\pi i t})^m e^{2\pi i k t}\, dt = \int_0^1 f(e^{2\pi i t})^n \bar{f}(e^{2\pi i t})^m\, dt.$$

According to (7), (9) and (5), this can be written in the form

$$\sum_{k=-\infty}^{\infty} \omega_{nk}(f)\bar{\omega}_{mk}(f) = \int_0^1 e^{2\pi i (n-m) x}\, d\phi(e^{2\pi i x}),$$

which, by the definitions (11), (10) and (2), is equivalent to

$$(22) \qquad \Omega(f)\Omega^*(f) = \Lambda(\phi).$$

This proves (II).

(III) is implied by (I) and (II), since, according to Hellinger and Toeplitz, a matrix A is bounded if and only if the matrix AA^* exists and is bounded.

Since (20) is necessary and sufficient for a non-singular matrix (15), and since (21) was shown to hold whenever the matrix $\Lambda(\phi)$ is bounded, (IV) follows from the definition, (14), of a function of bounded distortion.

Finally, (V) follows from (II), (III) and (IV), since, according to Toeplitz, a bounded matrix A is non-singular if and only if both matrices AA^*, A^*A are non-singular.

4. If f has an absolutely convergent Fourier series, then

$$\sum_{m=-\infty}^{\infty} |\omega_{nm}(f)| < \infty \tag{23}$$

for every n. In fact, if $f(t)$ has an absolutely convergent Fourier series, then (23) holds, by (6), for $n = 1$. Hence, the case $n = \pm 1$ of (8) shows that (23) holds for $n = -1$ also. Since the product of two functions possessing absolutely convergent Fourier series possesses an absolutely convergent Fourier series, it follows from (7) that (23) holds for every $n \gtrless 0$ (and, of course, for $n = 0$ also).

Let f_1, f_2 denote measurable functions f satisfying (5).

(VI) *If $f_1 = f_1(e^{2\pi i t})$ has an absolutely convergent Fourier series, the matrix product $\Omega(f_1)\Omega(f_2)$ exists for every $f_2 = f_2(e^{2\pi i t})$ and is identical with $\Omega(f_1(f_2))$, where $f_1(f_2) = f_1(f_2(e^{2\pi i t}))$.*

It is obvious from (III) that the assumption of (VI) does not imply the boundedness of any of the three matrices $\Omega(f_1)$, $\Omega(f_2)$, $\Omega(f_1(f_2))$.

Since (23) is satisfied by $f = f_1$, one sees from (6) that

$$\sum_{m=-\infty}^{\infty} \omega_{nm}(f_1)\,(e^{2\pi i t})^m = f_1(e^{2\pi i t})^n$$

for every t. Since $|f_2(e^{2\pi i t})| = 1$ for every t, it follows that

$$\sum_{m=-\infty}^{\infty} \omega_{nm}(f_1) f_2(e^{2\pi i t})^m = f_1(f_2(e^{2\pi i t}))^n.$$

On multiplying this relation by $e^{-2\pi i k t}$ and noting that (23), where $f = f_1$, insures the term-by-term integrability of the resulting series, one sees from the definition (7), that

$$\sum_{m=-\infty}^{\infty} \omega_{nm}(f_1)\,\omega_{mk}(f_2) = \omega_{nk}(f_1(f_2)).$$

According to (11), this is equivalent to

$$\Omega(f_1)\Omega(f_2) = \Omega(f_1(f_2)); \tag{24}$$

so that the proof of (VI) is complete.

If f is a measurable function satisfying (5), the same is true of the j-th iterate, $f^{(j)}$, where

$$(25) \qquad f^{(j+1)}(e^{2\pi it}) = f(f^{(j)}(e^{2\pi it})), \qquad j = 1, 2, \cdots ; f^{(1)} = f.$$

If $j = 0$, let

$$(26) \qquad f^{(0)}(e^{2\pi it}) = e(e^{2\pi it}), \text{ where } e(e^{2\pi it}) = e^{2\pi it},$$

it being understood that the first e in $e = e(e^{2\pi it})$ is meant as a function symbol f.

(VII) *If* $f(e^{2\pi it})$ *is of constant absolute value* 1 *and has an absolutely convergent Fourier series, then*

$$(27) \qquad \Omega(f^{(j)}) = \Omega^j(f)$$

for $j = 0, 1, 2, \cdots$.

It is understood that the matrix on the right of (17) denotes the j-th power of $\Omega(f)$. Thus (VII) is clear from (VI) and from the definition (25), if $j > 0$; while if $j = 0$, then (27) reduces, by (26), to the relation

$$(28) \qquad \Omega(e) = E, \qquad (E = \text{unit matrix}),$$

which is obvious from the definitions (11), (7).

If $f = f(e^{2\pi it})$ is a one-to-one transformation of the unit circle, $|z| = 1$, into itself, let $f^{(-1)} = f^{(-1)}(e^{2\pi it})$ denote the inverse mapping; so that

$$(29_1) \qquad f(f^{(-1)}(e^{2\pi it})) = e^{2\pi it}; \qquad (29_2) \qquad f^{(-1)}(f(e^{2\pi it})) = e^{2\pi it}$$

for every t.

(VIII) *If a function* $f = f(e^{2\pi it})$ *of constant absolute value* 1 *represents a one-to-one continuous transformation of the unit circle into itself and possesses an absolutely convergent Fourier series, then*

$$(30) \qquad \Omega(f)\Omega(f^{(-1)}) = E$$

(*the existence of the product on the left of* (30) *being a part of the statement*).

It is clear from (29_1) and (28) that (VIII) is the particular case $f_1 = f, f_2 = f^{(-1)}$ of (VI).

Since the assumptions of (VIII) and the assertion of (VI) are not symmetric in the two functions, it does not follow that (30) can be replaced by

$$(30 \text{ bis}) \qquad \Omega(f^{(-1)})\Omega(f) = E,$$

although not only (29_1) but also (29_2) is supposed in (VIII).

5. It turns out that this dissymmetry in (VIII) and the corresponding dissymmetry in (V) disappear if one combines the corresponding conditions.

THEOREM. *If a function* $f = f(e^{2\pi it})$ *of constant absolute value* 1 *represents a continuous one-to-one mapping of the unit circle into itself, the matrix* $\Omega(f)$ *is non-singular if and only if the function* $f(e^{2\pi it})$ *is of bounded distortion; in which case the reciprocal matrix,* $\Omega^{-1}(f)$, *is precisely* $\Omega(f^{(-1)})$.

Let $f = f(e^{2\pi it})$ be a one-to-one continuous mapping of the unit circle on itself. Then it is clear, from the remark made after (9), that

$$2\pi\phi(e^{2\pi it}) = \arg f^{(-1)}(e^{2\pi it}),$$

where ϕ denotes the distribution function, and $f^{(-1)}$ the inverse function, of the function f (mod 1). Since this correspondence is involutory, it follows that

$$2\pi\phi^{(-1)}(e^{2\pi it}) = \arg f(e^{2\pi it}),$$

where $\phi^{(-1)}$ denotes the distribution function of $f^{(-1)}$. Thus it is clear that all four functions f, $f^{(-1)}$, ϕ, $\phi^{(-1)}$ are of bounded distortion whenever one of them is. But if a function is of bounded distortion, then it satisfies a uniform Lipschitz condition and has, therefore, an absolutely convergent Fourier series, provided that it is periodic. Consequently, (VIII) is applicable not only to f but also to $f^{(-1)}$, if f is of bounded distortion. Furthermore, if f is of bounded distortion, then, since ϕ and $\phi^{(-1)}$ are of bounded distortion, ϕ and $\phi^{(-1)}$ satisfy uniform Lipschitz conditions; so that $\Omega(f)$ and $\Omega(f^{-1})$ are bounded, by (III). Hence, in order to complete the proof of the theorem, it is sufficient to ascertain that $\Omega(f)$ cannot be non-singular unless f is of bounded distortion. But this is implied by (V), since ϕ cannot be of bounded distortion unless f is of bounded distortion.

THE JOHNS HOPKINS UNIVERSITY.

SEMI-SIMPLE ALGEBRAS AND GENERALIZED DERIVATIONS.*

By G. Hochschild.

Introduction. Roughly speaking, a derivation of an algebra is the infinitesimal operation corresponding to an automorphism. This notion plays an important part in the study of the structure of Lie algebras and associative algebras, and its structural significance will be understood best by first recalling the connections between derivations and automorphisms in general.

The term "algebra" will be used to denote a linear set (not necessarily of finite dimension) over a field, in which an operation of "multiplication" is defined as usual, except that we do not require associativity. By a derivation of such an algebra we shall mean a linear mapping of the algebra into itself which possesses the formal properties of ordinary differentiation, i. e. a linear mapping $a \rightarrow D(a)$ of an algebra $\mathfrak{A}$ into itself such that for any $x, y \epsilon \mathfrak{A}$ we have $D(x \cdot y) = D(x) \cdot y + x \cdot D(y)$, where the dot indicates the operation of "multiplication" in $\mathfrak{A}$.[1]

A simple example is given by the ring of all polynomials (with real coefficients) in one variable. We may regard this ring as an algebra (in the above general sense) over the field of real numbers. The mapping which sends every polynomial $P(x)$ into its derivative $D_x(P)$ is clearly a derivation in the above sense. Consider now the operator T_t, depending on the real parameter t, which is defined by the formula

$$T_t = 1 + tD_x + (1/2!)t^2D_x^2 + \cdots = \exp(tD_x).$$

Then, if $P(x)$ is any polynomial, we clearly have $T_t(P(x)) = P(x+t)$, and the family of mappings $\{T_t\}$ constitutes a continuous one parameter group of automorphisms.

Let us now consider a more general situation: Let $\mathfrak{A}$ be an arbitrary algebra of finite dimension over the field of the real numbers. The set of all non singular linear transformations of $\mathfrak{A}$ into itself can be given the structure of a Lie group G. The corresponding Lie algebra, $\mathfrak{L}$, is the set of all linear transformations of $\mathfrak{A}$, commutation being defined by the formula $T_1 \circ T_2 = T_2T_1 - T_1T_2$. If T is any element of $\mathfrak{L}$, we obtain a one parameter sub-

---

* Received May 1, 1941.

[1] This definition and the elementary theory of derivations will be found in N. Jacobson, "Abstract derivation and Lie algebras," *Transactions of the American Mathematical Society*, vol. 42 (1937), p. 206.

group $g = \{g_t\}$ of G, whose tangent at the identity E of G is T, if we define $g_t = E + t \cdot T + (1/2!)t^2 \cdot T^2 + \cdots = \exp(t \cdot T)$. Since the condition that a transformation $g \in G$ be an automorphism can be expressed by a finite number of analytic equations in the coefficients of the matrix representing the element g, it is clear that the automorphisms of $\mathfrak{A}$ constitute a Lie-subgroup (i. e. a closed subgroup) H of G.

On the other hand, the derivations of $\mathfrak{A}$ constitute a subalgebra $\mathfrak{D}$ of $\mathfrak{L}$; for, if D_1, D_2 are any two derivations of $\mathfrak{A}$, so is $D_1 \circ D_2$. We shall show that $\mathfrak{D}$ is the subalgebra of $\mathfrak{L}$ which corresponds to the subgroup H of G.

Let $h = (h_t)$ be a one parameter subgroup of H which passes through the identify E for $t = 0$, and let $L \in \mathfrak{L}$ be its tangent at E, i. e. $L = \lim\limits_{t \to 0} (h_t - E)/t$. If x, y are arbitrary elements of $\mathfrak{A}$, we have $h_t(x \cdot y) = h_t(x) \cdot h_t(y)$, which is easily seen to imply that $L(x \cdot y) = L(x) \cdot y + x \cdot L(y)$, i. e. that L is a derivation, or $L \in \mathfrak{D}$.

Conversely, let $D \in \mathfrak{D}$; the corresponding one parameter subgroup of G is given by $g_t = \exp(t \cdot D)$, and a direct computation shows that $g_t(x \cdot y) = g_t(x) \cdot g(y)$, i. e. that $g_t \in H$. This proves our assertion.

Let $\mathfrak{A}$ be an associative algebra with a unity element. If c is any regular element of $\mathfrak{A}$, the automorphism $a \to c. a. c^{-1}$ is called an inner automorphism. Algebraically, the inner automorphisms constitute an invariant subgroup K of the group H of all automorphisms of $\mathfrak{A}$. Although we cannot say in general that K is also a topological (Lie-) subgroup of H, it is evident that we can define a topology in K such that it becomes a Lie group K^* which is locally isomorphic with the group K, regarded as a subspace of H. It follows that there corresponds to K a certain ideal $\mathfrak{J}$ of the Lie algebra $\mathfrak{D}$ of H. In fact, the elements of $\mathfrak{J}$ are the tangents at E of those one parameter subgroups of H which are contained in K, and if k_t is such a subgroup the corresponding element $I \in \mathfrak{J}$ is given by the equation $I = \lim\limits_{t \to 0} (k_t - E)/t$, so that we have $k_t(a) = a + t \cdot I(a) + o(t)$, for every element $a \in \mathfrak{A}$.

The group constituted by the regular elements of $\mathfrak{A}$ can be given the structure of a Lie group, and it can be shown that the one parameter subgroups, in a sufficiently small neighborhood of the identity e, are of the form $a_t = \exp(ta_0)$, where a_0 is an element of $\mathfrak{A}$. Furthermore, it is easily seen that we can find an analytic mapping $k_t \to a_t$ of any given one parameter subgroup of K onto a one parameter subgroup through e of the group of regular elements of $\mathfrak{A}$ such that $k_t(a) = a_t \cdot a \cdot a_t^{-1}$ for every $a \in \mathfrak{A}$. For sufficiently small t we have therefore $k_t(a) = \exp(ta_0) \cdot a \cdot \exp(-ta_0) = a + t(a_0 \cdot a - a \cdot a_0) + o(t)$, and a comparison of this with the expression above shows that we have $I(a) = a_0 \cdot a - a \cdot a_0$. Derivations of this form are called inner derivations.

Suppose that I is any given inner derivation, i. e. $I(a) = a_0 \cdot a - a \cdot a_0$; the corresponding one parameter group of automorphisms is given by $k_t = \exp(tI)$. A straightforward computation shows that $k_t(a) = a_t \cdot a \cdot a_t^{-1}$, where $a_t = \exp(ta_0)$, i. e. the automorphisms corresponding to the inner derivations are the inner automorphisms. We have shown, therefore, that $\mathfrak{J}$ is the set of inner derivations of $\mathfrak{A}$.

We have a similar situation in the case of a Lie algebra. Let $\mathfrak{L}$ be any Lie algebra over the field of real numbers and let Γ be a connected and simply connected Lie group whose Lie algebra is $\mathfrak{L}$. An automorphism of Γ will induce an automorphism of $\mathfrak{L}$, and conversely. An automorphism of $\mathfrak{L}$ will be called an inner automorphism if it corresponds to an inner automorphism of Γ. As in the case of an associative algebra, it can be shown that to the invariant subgroup of the group of automorphisms which consists of the inner automorphisms there corresponds an ideal of the Lie algebra of the derivations of $\mathfrak{L}$, whose elements will be called inner derivations. A derivation D of the Lie algebra $\mathfrak{L}$ is an inner derivation if and only if there exists an element $l_0 \in \mathfrak{L}$ such that $D(l) = l \circ l_0$, for every $l \in \mathfrak{L}$. In the case of a general groundfield this is taken as the definition of an inner derivation.

We shall be concerned with the study of the behavior of Lie algebras and associative algebras with regard to linear mappings of the derivational type which map a given algebra $\mathfrak{A}$, not necessarily into itself, but, more generally, into some algebra $\mathfrak{B}$ containing $\mathfrak{A}$ as a subalgebra. These "generalized derivations," as we should expect from the above considerations, are found to be significant for the structure of an algebra. In fact, we shall obtain a characterization of semi-simple Lie algebras and semi-simple associative algebras (over a non-modular field) in terms of their generalized derivations. In this respect there is a very close analogy between Lie algebras and associative algebras, which we shall exhibit by treating the two cases side by side.

I have found it necessary to include proofs of a number of known results on representations of semi-simple Lie algebras, some of which (like the proofs of Whitehead's lemmas) cannot be found in the literature. For these proofs, which served me as a guide in the present investigation, I am indebted to Professor C. Chevalley of Princeton University.

1. Representations.

For the sake of future reference we briefly review some fundamental results of the theory of representations of associative and Lie algebras.

Definition 1. 1. A representation is said to be completely decomposable if the corresponding representation space is the direct sum of irreducible invariant subspaces.

Proposition 1.1. A representation, with the representation space $\mathfrak{P}$, is completely decomposable if and only if for every invariant subspace $\mathfrak{Q}$ of $\mathfrak{P}$ there exists another invariant subspace $\mathfrak{Q}'$ such that $\mathfrak{P}$ is the direct sum of $\mathfrak{Q}$ and $\mathfrak{Q}'$.

The proof of this proposition is very easy and may be omitted here.

THEOREM 1.1. *The necessary and sufficient condition for an associative algebra to be semi-simple is that every representation be completely decomposable.*

This result is classical, and no proof need be given here.

THEOREM 1.2. *A sufficient condition for a Lie algebra to be semi-simple is that every representation be completely decomposable.*

Proof. We must show that a Lie algebra $\mathfrak{L}$ satisfying this condition cannot possess a (non zero) solvable ideal.

Let $\mathfrak{H}$ be any solvable ideal of $\mathfrak{L}$. Since in particular the adjoint representation of $\mathfrak{L}$ is completely decomposable we can find another ideal $\mathfrak{H}^*$ of $\mathfrak{L}$ such that $\mathfrak{L}$ is the direct sum of $\mathfrak{H}$ and $\mathfrak{H}^*$. Hence it is clear that every representation of $\mathfrak{H}$ can be extended to a representation of $\mathfrak{L}$ by representing the elements of $\mathfrak{H}^*$ by zero matrices. It follows that $\mathfrak{H}$ still has the property that all representations are completely decomposable. Hence it will suffice to show that every solvable Lie algebra $\mathfrak{H}$ of dimension $d > 0$ possesses a representation which is not completely decomposable.

Now, if $\mathfrak{H}$ is solvable it contains an ideal $\mathfrak{H}^*$ of dimension $d - 1$. Let h_1 be an element of $\mathfrak{H}$ which does not belong to $\mathfrak{H}^*$; then every element $h \in \mathfrak{H}$ can be written in the form $h = \alpha(h)h_1 + h^*$, where $\alpha(h)$ is a linear function on $\mathfrak{H}$ (with values in the groundfield) and $h^* \in \mathfrak{H}^*$. Let $\mathfrak{P}$ be a linear space of dimension 2 over the groundfield, and let the vectors $\vec{e}_1$, $\vec{e}_2$ constitute a basis for $\mathfrak{P}$.

If we write $P_h(\vec{e}_1) = \alpha(h)\vec{e}_1$; $P_h(\vec{e}_2) = \alpha(h)(\vec{e}_1 + \vec{e}_2)$, the mapping $h \to P_h$ is clearly a representation of $\mathfrak{H}$ in which $\mathfrak{H}^*$ is annulled. The subspace spanned by $\vec{e}_1$ is invariant, but it is clearly the only invariant subspace of dimension 1. It follows that the representation P is not completely decomposable. This completes the proof.

THEOREM 1.3. (Cartan). *Let $\mathfrak{L}$ be a Lie algebra over a field of characteristic zero. If there exists a faithful representation $l \to P_l$ of $\mathfrak{L}$, such that $\mathrm{Sp}(P_l^2) = 0$ for every $l \in \mathfrak{L}$, (where $\mathrm{Sp}(A)$ denotes the trace of the matrix A), then $\mathfrak{L}$ is solvable.*

THEOREM 1.4. (Cartan). *Let $\mathfrak{K}$ be a semi-simple Lie algebra over a field of characteristic zero. Let $k \rightarrow P_k$ be a faithful representation of $\mathfrak{K}$, and suppose that $k_1, k_2, \cdots, k_n$ constitute a basis for $\mathfrak{K}$. Then the determinant $|\operatorname{Sp}(P_{k_i}P_{k_j})| \neq 0$.*

Both these results are well known.

THEOREM 1.5. *If $\mathfrak{K}$ is a semi-simple Lie algebra over a field of characteristic 0, the adjoint representation A of $\mathfrak{K}$ is completely decomposable.*

Proof. Let $\mathfrak{H}$ be any ideal of $\mathfrak{K}$. We have to show that there exists another ideal $\mathfrak{H}^*$ of $\mathfrak{K}$ such that $\mathfrak{K}$ is the direct sum of $\mathfrak{H}$ and $\mathfrak{H}^*$. Let $\mathfrak{H}^*$ be the set of all elements h^* of $\mathfrak{K}$ such that $\operatorname{Sp}(A_h A_{h^*}) = 0$ for all $h \in \mathfrak{H}$. $\mathfrak{H}^*$ is an ideal of $\mathfrak{K}$, and since the adjoint representation of a semi-simple Lie algebra is faithful we may conclude from Theorem 1.3 that $\mathfrak{H}^* \cap \mathfrak{H}$ is a solvable ideal. Hence we must have $\mathfrak{H}^* \cap \mathfrak{H} = \{0\}$. A dimension argument now shows that $\mathfrak{K}$ is the direct sum of $\mathfrak{H}$ and $\mathfrak{H}^*$.

THEOREM 1.6. *Let $\mathfrak{A}$ be a semi-simple associative algebra over a field of characteristic 0. Let $(a_1, a_2, \cdots, a_n)$ be a basis of $\mathfrak{A}$, and suppose that $P_{a_1}, P_{a_2}, \cdots, P_{a_n}$ are the matrices corresponding to the elements a_i of this basis in a faithful representation P of $\mathfrak{A}$. Then $|\operatorname{Sp}(P_{a_i}P_{a_j})| \neq 0$.*

Proof. Let us write $\mathfrak{A}$ as a direct sum of two sided simple ideals, say $\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 + \cdots + \mathfrak{A}_k$. We have then k sets of indices, $I_1, I_2, \cdots, I_k$ such that $a_i \in \mathfrak{A}_\nu$ for $i \in I_\nu$. Now $P_{a_i}P_{a_j} = 0$ if i and j do not belong to the same set I_ν, whence it is clear that

$$|\operatorname{Sp}(P_{a_i}P_{a_j})|_{(i,j=1,2\ldots n)} = \prod_{\nu=1}^{k} |\operatorname{Sp}(P_{a_i}P_{a_j})|_{(i,j\in I_\nu)}.$$

Every irreducible representation of an algebra is equivalent to an irreducible representation contained in the adjoint representation, and a simple algebra possesses only one irreducible representation (to within equivalence). Hence, if we denote the matrices in the adjoint representation by A, we have, for $i, j \in I_\nu$, $\operatorname{Sp}(P_{a_i}P_{a_j}) = r_\nu \cdot \operatorname{Sp}(A_{a_i}A_{a_j})$, where r_ν is a positive rational number depending only on ν. It follows that $|\operatorname{Sp}(P_{a_i}P_{a_j})|$ is a (non-zero) multiple of $|\operatorname{Sp}(A_{a_i}A_{a_j})|$, which is known to be different from zero for a semi-simple algebra. (Cf. Deuring, "Algebren," *Ergebnisse der Mathematik*, vol. 4, part 1, pp. 33-34).

THEOREM 1.7. (Casimir). *Let $\mathfrak{L}$ be a semi-simple Lie algebra over a field of characteristic zero. Let P be any (non null) representation of $\mathfrak{L}$, and let $\mathfrak{H}^*$ be the kernel of this representation* (i. e. *the ideal of $\mathfrak{L}$ which is annulled* *in the representation* P). *Let* $\mathfrak{H}$ *be another ideal such that* $\mathfrak{L}$ *is the direct sum of* $\mathfrak{H}$ *and* $\mathfrak{H}^*$; ($\mathfrak{H}$ *exists, by Theorem* 1.5). *If* $(h_1, h_2, \cdots, h_k)$ *is any basis of* $\mathfrak{H}$ *there exists a dual basis* $(\bar{h}_1 \cdot h_2, \cdots, \bar{h}_k)$ *of* $\mathfrak{H}$ *such that* $\mathrm{Sp}(P_{h_i} P_{\bar{h}_j}) = \delta_{ij}$, *where* δ_{ij} *is the Kronecker symbol* ($= 0$ *for* $i \neq j$, *and* $= 1$ *for* $i = j$).

Then, if l *is any element of* $\mathfrak{L}$, *and* $h_i \circ l = \sum_{j=1}^{k} \gamma_{ij} h_j$, *we have* $l \circ \bar{h}_i = \sum_{j=1}^{k} \gamma_{ji} \bar{h}_j$. *If we define a transformation* Γ *by the equation* $\Gamma = \sum_{i=1}^{k} P_{h_i} P_{\bar{h}_i}$, *we have* $\Gamma P_l = P_l \Gamma$ *for every* $l \in \mathfrak{L}$.

Proof. The existence of the dual basis $(\bar{h}_i)$ follows immediately from Theorem 1.4 since P gives a faithful representation of $\mathfrak{H}$. If $h_i \circ l = \sum_{j=1}^{k} \gamma_{ij} h_j$ and

$l \circ \bar{h}_j = \sum_{i=1}^{k} \epsilon_{ij} \bar{h}_i$, we have $\gamma_{ij} = \mathrm{Sp}(P_{h_i \circ l} P_{\bar{h}_j}) = \mathrm{Sp}(P_{h_i} P_{l \circ \bar{h}_j}) = \epsilon_{ij}$. Finally,

$$\Gamma P_l - P_l \Gamma = \sum_{i=1}^{k} (P_{h_i} P_{\bar{h}_i} P_l - P_l P_{h_i} P_{\bar{h}_i}) = \sum_{i=1}^{k} (P_{h_i} P_{l \circ \bar{h}_i} - P_{h_i \circ l} P_{\bar{h}_i})$$
$$= \sum_{i=1}^{k} \sum_{j=1}^{k} (\gamma_{ji} P_{h_i} P_{\bar{h}_j} - \gamma_{ij} P_{h_j} P_{\bar{h}_i}) = 0.$$

THEOREM 1.8. *Let* $\mathfrak{A}$ *be a semi-simple associative algebra over a field of characteristic zero. Let* P *be any (non null) representation of* $\mathfrak{A}$, *and let* $\mathfrak{H}^*$ *be the kernel of this representation.* $\mathfrak{H}^*$ *is a two sided ideal of* $\mathfrak{A}$, *and there exists another two sided ideal* $\mathfrak{H}$ *such that* $\mathfrak{A}$ *is the direct sum of* $\mathfrak{H}$ *and* $\mathfrak{H}^*$. *To any basis* $(h_1, h_2, \cdots, h_k)$ *of* $\mathfrak{H}$ *we can find a dual basis* $(\bar{h}_1, \bar{h}_2, \cdots, \bar{h}_k)$ *such that* $\mathrm{Sp}(P_{h_i} P_{\bar{h}_j}) = \delta_{ij}$. *If* a *is any element of* $\mathfrak{A}$, *and* $h_i \cdot a = \sum_{j=1}^{k} \gamma_{ij} h_j$, *we have* $a \cdot \bar{h}_i = \sum_{j=1}^{k} \gamma_{ji} \bar{h}_j$. *The transformation* $\Gamma = P_h$, *where* $h = \sum \bar{h}_i h_i$, *commutes with all the transformations of the representation* P.

Proof. The existence of the dual basis follows from Theorem 1.6. The remaining relations are easily verified as in the proof of Theorem 1.7.

2. Whitehead's lemmas and their analogues for associative algebras.

It is known that every representation of a semi-simple Lie algebra over a field of characteristic zero is completely decomposable. This result follows from a lemma of Whitehead's which can be stated as follows:

THEOREM 2.1. *Let* $\mathfrak{L}$ *be a semi-simple Lie algebra over a field of characteristic zero. Let* $\mathfrak{P}$ *be a representation module of* $\mathfrak{L}$, *and denote the transform* *of a vector* $\vec{e} \in \mathfrak{P}$ *by the transformation corresponding to an element* $l \in \mathfrak{L}$ *by* $l \cdot \vec{e}$. *Suppose we have a linear mapping* $l \rightarrow \vec{e}(l)$ *of* $\mathfrak{L}$ *into* $\mathfrak{P}$ *such that* $\vec{e}(l \circ m) = m \cdot \vec{e}(l) - l \cdot \vec{e}(m)$ *for all* $l, m \in \mathfrak{L}$. *Then there exists a fixed vector* $\vec{e}_0$ *in* $\mathfrak{P}$ *such that* $\vec{e}(l) = l \cdot \vec{e}_0$ *for every* $l \in \mathfrak{L}$.

Proof. If $\mathfrak{P}$ is annulled by $\mathfrak{L}$ we may take $\vec{e}_0 = 0$, and there is nothing to prove. If not, we employ the notations and results of Theorem 1.7 and define a vector $\vec{f}_0 = \sum_{i=1}^{k} h_i \cdot \vec{e}(\bar{h}_i)$. We have then

$$\begin{aligned} l \cdot \vec{f}_0 &= \sum_{i=1}^{k} l \cdot (h_i \cdot \vec{e}(\bar{h}_i)) = \sum_{i=1}^{k} [(h_i \circ l) \cdot \vec{e}(\bar{h}_i) + h_i \cdot (l \cdot \vec{e}(\bar{h}_i))] \\ &= \sum_{i=1}^{k} [h_i \cdot \vec{e}(l \circ \bar{h}_i) + h_i \cdot (l \cdot \vec{e}(\bar{h}_i))] = \sum_{i=1}^{k} h_i \cdot (\bar{h}_i \cdot \vec{e}(l)) = \Gamma\{\vec{e}(l)\}. \end{aligned}$$

We are now in a position to prove our theorem by an induction on the dimension d of the representation space $\mathfrak{P}$. By the remark made at the outset we may suppose that $d > 0$, and that the theorem holds for every representation space of dimension $< d$. If our representation is not the null representation (which we may suppose), we have $\mathrm{Sp}(\Gamma) = k \neq 0$, since the groundfield is of characteristic 0. Hence $\Gamma \neq 0$; therefore, if $\mathfrak{Q}$ is the linear subspace formed by the vectors $\vec{f} \in \mathfrak{P}$ such that $\Gamma(\vec{f}) = 0$, we have $\mathfrak{Q} \neq \mathfrak{P}$. If $\mathfrak{Q} = \{0\}$, Γ is a linear isomorphism of $\mathfrak{P}$ and therefore possesses an inverse Γ^{-1} which also commutes with all the transformations of our representation. Hence, if we put $\vec{e}_0 = \Gamma^{-1}(\vec{f}_0)$, we have $l \cdot \vec{e}_0 = l \cdot \{\Gamma^{-1}(\vec{f}_0)\} = \Gamma^{-1}(l \cdot \vec{f}_0) = \vec{e}(l)$, as required by the theorem. We may therefore suppose that $\mathfrak{Q} \neq \{0\}$. Now $\mathfrak{Q}$ is an invariant subspace of $\mathfrak{P}$ since $\Gamma(\vec{f}) = 0$ implies that $\Gamma(l \cdot \vec{f}) = l \cdot \Gamma(\vec{f}) = 0$. Hence our representation in $\mathfrak{P}$ induces a representation in $\mathfrak{P}/\mathfrak{Q}$, and, if $\vec{E}(l)$ denotes the coset, modulo $\mathfrak{Q}$, of $\vec{e}(l)$, the mapping $l \rightarrow \vec{E}(l)$ of $\mathfrak{L}$ into $\mathfrak{Q}/\mathfrak{P}$ still satisfies the conditions of our theorem. Since $\mathfrak{P}/\mathfrak{Q}$ is of dimension $< d$, it follows from our inductive hypothesis that there exists an element $\vec{E}_1 \in \mathfrak{P}/\mathfrak{Q}$ such that $\vec{E}(l) = l \cdot \vec{E}_1$ for every $l \in \mathfrak{L}$. Let $\vec{e}_1$ be any vector in this coset $\vec{E}_1$; if we put $\vec{f}(l) = \vec{e}(l) - l \cdot \vec{e}_1$, the mapping $l \rightarrow \vec{f}(l)$ is a linear mapping of $\mathfrak{L}$ into $\mathfrak{Q}$ such that $\vec{f}(l \circ m) = m \cdot \vec{f}(l) - l \cdot \vec{f}(m)$. Since $\mathfrak{Q}$ is also of dimension $< d$ there exists a vector $\vec{e}_2 \in \mathfrak{Q}$ such that $\vec{f}(l) = l \cdot \vec{e}_2$ for all $l \in \mathfrak{L}$. The vector $\vec{e}_0 = \vec{e}_1 + \vec{e}_2$ will be the required vector, and our theorem is proved.

Before stating the analogue of Theorem 2. 1 for associative algebras we make the following definition:

Definition 2. 1. Let $\mathfrak{A}$ be an associative algebra over a field $\mathfrak{F}$. Let $\mathfrak{P}$ be a linear set over $\mathfrak{F}$ which is at the same time an $\mathfrak{A}$-right and an $\mathfrak{A}$-left module, the operations being such that (apart from the usual associativity requirements for one sided modules) we have $a \cdot (\vec{e} \cdot b) = (a \cdot \vec{e}) \cdot b$ for all $a, b \in \mathfrak{A}$ and every $\vec{e} \in \mathfrak{P}$. Then $\mathfrak{P}$ is called a two sided $\mathfrak{A}$-module.

THEOREM 2. 2. *Let $\mathfrak{A}$ be a semi-simple associative algebra over a field $\mathfrak{F}$ of characteristic* 0. *Let $\mathfrak{P}$ be a two sided $\mathfrak{A}$-module over $\mathfrak{F}$, and suppose we have a linear mapping $a \rightarrow \vec{e}(a)$ of $\mathfrak{A}$ into $\mathfrak{P}$ such that $\vec{e}(a \cdot b) = a \cdot \vec{e}(b) + \vec{e}(a) \cdot b$. Then there exists $\vec{e}_0 \in \mathfrak{P}$ such that $\vec{e}(a) = a \cdot \vec{e}_0 - \vec{e}_0 \cdot a$.*

Proof. In the notation of Theorem 1. 8 we have:

$$h \cdot \vec{e}(a) = \sum_{i=1}^{k} \bar{h}_i h_i \cdot \vec{e}(a) = \sum_{i=1}^{k} \bar{h}_i \cdot [\vec{e}(h_i a) - \vec{e}(h_i) \cdot a]$$
$$= \sum_{i=1}^{k} a\bar{h}_i \cdot \vec{e}(h_i) - \sum_{i=1}^{k} \bar{h}_i \cdot \vec{e}(h_i) \cdot a = a \cdot \vec{f}_0 - \vec{f}_0 \cdot a,$$

where $\vec{f}_0 = \sum_{i=1}^{k} \bar{h}_i \cdot \vec{e}(h_i)$.

The proof now runs like that of Theorem 2. 1; we may assume that the left representation in $\mathfrak{P}$ is not null, for otherwise we could operate in the same manner from the right, while—if both the left and the right representations were null—we could take $\vec{e}_0 = 0$. The subspace $\mathfrak{Q}$, defined as before, is invariant for both the right and the left representation, and the induction goes exactly as before.

An interesting application of Theorem 2. 2 is the following: Let P_n be the algebra of all matrices of order n over a field P of characteristic 0. Let $\mathfrak{A}$ be a semi-simple subalgebra and U an arbitrary matrix of P_n. The elements of the form $UA - AU$, where $A \in \mathfrak{A}$, together with all products of such elements by elements of $\mathfrak{A}$, span a certain linear subspace of P_n, which we may regard as a two sided $\mathfrak{A}$-module; $\mathfrak{M}$, say. It is easily verified that the mapping $A \rightarrow UA - AU$ of $\mathfrak{A}$ into $\mathfrak{M}$ satisfies the conditions of Theorem 2. 2. It follows that there exists a matrix $M \in \mathfrak{M}$ such that $U - M$ commutes with all the matrices of $\mathfrak{A}$.

THEOREM 2. 3. *If an associative algebra $\mathfrak{A}$ possesses the property of Theorem 2. 2 then every representation of $\mathfrak{A}$ is completely decomposable.*

Proof. Let $\mathfrak{P}$ be the representation space of a representation P of $\mathfrak{A}$. If $\mathfrak{Q}$ is any invariant subspace of $\mathfrak{P}$ we have to show that there exists another invariant subspace $\mathfrak{R}$ such that $\mathfrak{P} = \mathfrak{Q} + \mathfrak{R}$. Let us choose a linear subspace $\mathfrak{T}$ (not necessarily invariant, of course) such that $\mathfrak{P} = \mathfrak{Q} + \mathfrak{T}$. Let P_a be the transformation of $\mathfrak{P}$ which corresponds to the element a of $\mathfrak{A}$. Let $\vec{t}$ denote an arbitrary vector in $\mathfrak{T}$. According to the decomposition $\mathfrak{P} = \mathfrak{Q} + \mathfrak{T}$ we write: $P_a(\vec{t}) = Q_a(\vec{t}) + T_a(\vec{t})$, where $Q_a(\vec{t}) \in \mathfrak{Q}$, $T_a(\vec{t}) \in \mathfrak{T}$. It is clear that the mapping $\vec{t} \rightarrow Q_a(\vec{t})$ is a linear mapping of $\mathfrak{T}$ into $\mathfrak{Q}$ and the mapping $\vec{t} \rightarrow T_a(\vec{t})$ a linear mapping of $\mathfrak{T}$ into itself. From the fact that P is a representation we obtain the following relations:

$$\begin{aligned} P_{ab}(\vec{t}) &= Q_{ab}(\vec{t}) + T_{ab}(\vec{t}) = P_aP_b(\vec{t}) = P_a(Q_b(\vec{t}) + T_b(\vec{t})) \\ &= (P_aQ_b(\vec{t}) + Q_aT_b(\vec{t})) + T_aT_b(\vec{t}). \end{aligned}$$

Since the sum $\mathfrak{Q} + \mathfrak{T}$ is direct it follows that we must have

$$Q_{ab} = P_aQ_b + Q_aT_b; \qquad T_{ab} = T_aT_b.$$

The set of all linear mappings Q of $\mathfrak{T}$ into $\mathfrak{Q}$ forms a linear space over our groundfield: we denote this linear space by $\mathfrak{S}$. Now we may regard $\mathfrak{S}$ as a two-sided $\mathfrak{A}$-module if we define the operations by elements $a \in \mathfrak{A}$ on elements $Q \in \mathfrak{S}$ by the formulae: $a \cdot Q = P_aQ$, $Q \cdot a = QT_a$. Consider the linear mapping $a \rightarrow Q_a$ of $\mathfrak{A}$ into $\mathfrak{S}$. We have $Q_{ab} = P_aQ_b + Q_aT_b = a \cdot Q_b + Q_a \cdot b$. Now, since $\mathfrak{A}$ possesses the property of Theorem 2. 2, there exists an element $Q_0 \in \mathfrak{S}$ such that $Q_a = Q_0T_a - P_aQ_0$.

Consider now the mapping $\vec{t} \rightarrow \vec{t} + Q_0(\vec{t})$ of $\mathfrak{T}$ into $\mathfrak{P}$. This maps $\mathfrak{T}$ onto some linear subspace $\mathfrak{R}$ of $\mathfrak{P}$. We have $P_a(\vec{t} + Q_0(\vec{t})) = P_a(\vec{t}) + P_aQ_0(\vec{t}) = P_a(\vec{t}) + Q_0T_a(\vec{t}) - Q_a(\vec{t}) = T_a(\vec{t}) + Q_0T_a(\vec{t})$, which shows that $\mathfrak{R}$ is an invariant subspace of $\mathfrak{P}$. Furthermore, $\vec{t} + Q_0(\vec{t}) = 0$ only if $\vec{t} = 0$, so that our mapping $\vec{t} \rightarrow \vec{t} + Q_0(\vec{t})$ is a linear isomorphism, whence we have that dim. $\mathfrak{R} =$ dim. $\mathfrak{T}$. Finally, if $\vec{e} \in \mathfrak{Q} \cap \mathfrak{R}$, i. e. if $\vec{e} \in \mathfrak{Q}$, and $\vec{e} = \vec{t} + Q_0(\vec{t})$ for some $\vec{t} \in \mathfrak{T}$, we must have $\vec{t} = 0$, and hence also $\vec{e} = 0$, whence $\mathfrak{Q} \cap \mathfrak{R} = \{0\}$. It follows that $\mathfrak{P}$ is the direct sum of $\mathfrak{Q}$ and $\mathfrak{R}$, and our theorem is proved.

The proof of Theorem 2. 3 we have just given is the precise analogue of a known proof for Theorem 2. 4 below:

THEOREM 2. 4. *If a Lie algebra $\mathfrak{L}$ possesses the property of Theorem 2. 1 then every representation of $\mathfrak{L}$ is completely decomposable.*

By Theorems 2. 1, 2. 4, and 1. 2 we have now:

THEOREM 2. 5. *The necessary and sufficient condition for a Lie algebra* $\mathfrak{L}$ *over a field of characteristic zero to be semi-simple is that every representation of* $\mathfrak{L}$ *be completely decomposable.*

We shall pursue the analogy between Lie algebras and associative algebras a little further: another lemma of Whitehead, which is used in the proof of Levi's theorem on the decomposition of a Lie algebra into the direct sum of its maximal solvable ideal and a semi-simple subalgebra, can be stated as follows:

THEOREM 2. 6. *Let* $\mathfrak{P}$ *be a representation space of a semi-simple Lie algebra* $\mathfrak{L}$ *over a field of characteristic* 0. *Suppose we have a bilinear mapping* $(x, y) \to \vec{f}(x, y)$ *of* $\mathfrak{L} \times \mathfrak{L}$ *into* $\mathfrak{P}$ *such that*

1) $\vec{f}(x, y) + \vec{f}(y, x) = 0,$

2) $x \cdot \vec{f}(y, z) + y \cdot \vec{f}(z, x) + z \cdot \vec{f}(x, y) = \vec{f}(y \circ x, z) + \vec{f}(z \circ y, x) + \vec{f}(x \circ z, y).$

Then there exists a linear mapping $x \to \vec{e}(x)$ *of* $\mathfrak{L}$ *into* $\mathfrak{P}$ *such that* $\vec{f}(x, y) = x \cdot \vec{e}(y) - y \cdot \vec{e}(x) + \vec{e}(x \circ y).$

We shall not write out the proof of this theorem. It is very similar to that of Theorem 2. 7 below and requires the same methods we have employed already in the proof of Theorem 2. 1.

There is an interesting analogue of this theorem for associative algebras; it is as follows:

THEOREM 2. 7. *Let* $\mathfrak{A}$ *be a semi-simple associative algebra over a field of characteristic* 0. *Let* $\mathfrak{P}$ *be a two-sided* $\mathfrak{A}$*-module and suppose we have a bilinear mapping* $(a, b) \to \vec{f}(a, b)$ *of* $\mathfrak{A} \times \mathfrak{A}$ *into* $\mathfrak{P}$ *such that* $a \cdot \vec{f}(b, c) + \vec{f}(a, bc) = \vec{f}(a, b) \cdot c + \vec{f}(ab, c)$. *Then there exists a linear mapping* $a \to \vec{e}(a)$ *of* $\mathfrak{A}$ *into* $\mathfrak{P}$ *such that* $\vec{f}(a, b) = a \cdot \vec{e}(b) + \vec{e}(a) \cdot b - \vec{e}(ab).$

Proof. In the notation of Theorem 1. 8 we have: $h \cdot \vec{f}(a, b) = \sum_{i=1}^{k} \bar{h}_i h_i \cdot \vec{f}(a, b) = \sum_{i=1}^{k} \bar{h}_i \cdot (\vec{f}(h_i, a) \cdot b + \vec{f}(h_i a, b) - \vec{f}(h_i, ab)) = \vec{e}_1(a) \cdot b - \vec{e}_1(ab) + \sum_{i=1}^{k} \bar{h}_i \cdot \vec{f}(h_i a, b)$, where we have put $\vec{e}_1(a) = \sum_{i=1}^{k} \bar{h}_i \cdot \vec{f}(h_i, a)$. Now, as a consequence of the properties of the bases (h_i) and $(\bar{h}_i)$, we have $\sum_{i=1}^{k} \bar{h}_i \cdot \vec{f}(h_i a, b) = \sum_{i=1}^{k} a\bar{h}_i \cdot \vec{f}(h_i, b) = a \cdot \vec{e}_1(b)$. Hence $h \cdot \vec{f}(a, b) = a \cdot \vec{e}_1(b) + \vec{e}_1(a) \cdot b - \vec{e}_1(ab)$.

Let now $\mathfrak{Q}$ be the linear subspace of $\mathfrak{P}$ formed by those elements $\vec{g}$ for which $h \cdot \vec{g} = 0$. Suppose first that $\mathfrak{Q} = \{0\}$. Then the mapping $\vec{e} \rightarrow h \cdot \vec{e} = \Gamma(\vec{e})$ possesses an inverse Γ^{-1}, which commutes with all the elements of $\mathfrak{A}$, regarded as operators on $\mathfrak{P}$, and we may then take $\vec{e}(a) = \Gamma^{-1}\{\vec{e}_1(a)\}$.

Next, let us suppose that both the left representation and the right representation are null. Our condition on the mapping $\vec{f}$ then becomes: $\vec{f}(a, bc) = \vec{f}(ab, c)$. Let $(\mathfrak{A}, \mathfrak{P})$ denote the direct sum of the linear spaces $\mathfrak{A}$ and $\mathfrak{P}$. To every element a of $\mathfrak{A}$ we have a transformation of $(\mathfrak{A}, \mathfrak{P})$ which is defined by the formula $(b, \vec{e}) \rightarrow a \cdot (b, \vec{e}) = (ab, \vec{f}(a, b))$; it is easily verified that this defines a representation of $\mathfrak{A}$ with $(\mathfrak{A}, \mathfrak{P})$ as representation space. The subspace $(0, \mathfrak{P})$ is clearly an invariant subspace. Since every representation of is completely decomposable there must exist another invariant subspace $\mathfrak{Q}^*$ of $(\mathfrak{A}, \mathfrak{P})$ such that $(\mathfrak{A}, \mathfrak{P})$ is the direct sum of $(0, \mathfrak{P})$ and $\mathfrak{Q}^*$. Let us write, according to this decomposition, $(a, 0) = (0, \vec{e}(a)) + q^*(a)$. We have then

$$b \cdot (a, 0) = b \cdot (0, \vec{e}(a)) + b \cdot q^*(a), \text{ i. e. } (ba, \vec{f}(b, a)) = b \cdot q^*(a) \in \mathfrak{Q}^*.$$

On the other hand, we have

$$(ba, \vec{f}(b, a)) = (ba, 0) + (0, \vec{f}(b, a)) = (0, \vec{e}(ba)) + q^*(ba) + (0, \vec{f}(b, a)).$$

It follows that $(0, \vec{e}(ba)) + (0, \vec{f}(b, a)) = 0$; the mapping $a \rightarrow -\vec{e}(a)$ is a linear mapping of $\mathfrak{A}$ into $\mathfrak{P}$ such that $\vec{f}(a, b) = -\vec{e}(ab)$, and our theorem is verified for this case. The theorem is trivial if $\mathfrak{P}$ is of dimension 0, and we can now proceed by induction as in the proof of Theorem 2.1, since we shall have $\mathfrak{Q} \neq \mathfrak{P}$ for either the left representation or for the right representation, unless we have the case just disposed of. As an application of Theorem 2.4 we give a proof of the well-known result that every associative algebra over a field of characteristic zero can be represented as the direct sum of its radical and a semi-simple subalgebra.[2]

---

[2] J. H. C. Whitehead has communicated a proof of this result to N. Jacobson, which utilizes the same ideas as the proof given here.

Let $\mathfrak{A}$ be any associative algebra over a field of characteristic zero, and let $\mathfrak{N}$ denote the radical of $\mathfrak{A}$. We may clearly suppose that $\mathfrak{N} \neq \{0\}$. If we have $\mathfrak{N}^2 \neq \{0\}$, $\mathfrak{A}/\mathfrak{N}^2$ is of smaller dimension than $\mathfrak{A}$. Hence, if we suppose that the result has already been proved for all algebras of smaller dimension than $\mathfrak{A}$ (it is obvious for dimension one), we have that $\mathfrak{A}/\mathfrak{N}^2 = \mathfrak{S}^* + \mathfrak{A}^*_1$, where $\mathfrak{S}^*$ is the radical of $\mathfrak{A}/\mathfrak{N}^2$ and $\mathfrak{A}^*_1$ is a semi-simple subalgebra of $\mathfrak{A}/\mathfrak{N}^2$ which is isomorphic with $(\mathfrak{A}/\mathfrak{N}^2)/\mathfrak{S}^*$. But it is clear that $\mathfrak{S}^* = \mathfrak{N}/\mathfrak{N}^2$, and hence that $\mathfrak{A}^*_1$ is isomorphic with $\mathfrak{A}/\mathfrak{N}$. Now, there clearly exists a subalgebra $\mathfrak{A}_1$ of $\mathfrak{A}$ such that $\mathfrak{A}^*_1 = \mathfrak{A}_1/\mathfrak{N}^2$, and we have then $\mathfrak{A}/\mathfrak{N}^2 = \mathfrak{N}/\mathfrak{N}^2 + \mathfrak{A}_1/\mathfrak{N}^2$. Since $\mathfrak{N}/\mathfrak{N}^2 \neq \{0\}$, we have dim. $\mathfrak{A}_1 <$ dim. $\mathfrak{A}$. Hence there exists a subalgebra $\mathfrak{A}_2$ of $\mathfrak{A}_1$, i. e. also of $\mathfrak{A}$, which is isomorphic with $\mathfrak{A}_1/\mathfrak{N}^2$, and hence also with $\mathfrak{A}/\mathfrak{N}$, such that $\mathfrak{A}_1 = \mathfrak{N}^2 + \mathfrak{A}_2$, and we have $\mathfrak{A} = \mathfrak{N} + \mathfrak{A}_2$.

The only difficulty arises if $\mathfrak{N}^2 = \{0\}$. In this case we note that $\mathfrak{N}$ may be regarded as a two-sided $\mathfrak{A}/\mathfrak{N}$-module, the operation with elements of $\mathfrak{A}/\mathfrak{N}$ on $\mathfrak{N}$ being that of multiplication of the elements of $\mathfrak{N}$ by an arbitrary representant in $\mathfrak{A}$ of an element of $\mathfrak{A}/\mathfrak{N}$. (Since $\mathfrak{N}^2 = \{0\}$, all representants of a given coset, modulo $\mathfrak{N}$, will induce the same transformation of $\mathfrak{N}$.)

Let us choose a linear subspace $\mathfrak{S}$ of $\mathfrak{A}$ such that $\mathfrak{A} = \mathfrak{N} + \mathfrak{S}$. We have then a one to one correspondence $a \leftrightarrow a^*$ between the elements a of $\mathfrak{A}/\mathfrak{N}$ and the elements a^* of $\mathfrak{S}$. Set $\bar{f}(a, b) = (ab)^* - a^*b^*$. Then $\bar{f}(a, b) \in \mathfrak{N}$ for any $a, b \in \mathfrak{A}/\mathfrak{N}$ and gives a bi-linear mapping of $\mathfrak{A}/\mathfrak{N} \times \mathfrak{A}/\mathfrak{N}$ into $\mathfrak{N}$. Now, we have $\bar{f}(a, b) \cdot c + \bar{f}(ab, c) = ((ab)^* - a^*b^*) \cdot c + (abc)^* - (ab)^*c^*$, $a \cdot \bar{f}(b, c) + \bar{f}(ab, c) = a \cdot ((bc)^* - b^*c^*) + (abc)^* - a^*(bc)^*$. But if $a \in \mathfrak{A}/\mathfrak{N}$, $r \in \mathfrak{N}$, we have $a \cdot r = a^*r$. Hence we have $\bar{f}(a, b) \cdot c + \bar{f}(ab, c) = (abc)^* - a^*b^*c^* = a \cdot \bar{f}(b, c) + \bar{f}(a, bc)$, i. e. $\bar{f}$ satisfies the conditions of Theorem 2. 4. It follows that there exists a mapping $a \rightarrow \tilde{a}$ of $\mathfrak{A}/\mathfrak{N}$ into $\mathfrak{N}$, such that $\bar{f}(a, b) = a^*\tilde{b} + \tilde{a}b^* - \widetilde{ab}$, i. e. $(ab)^* - a^*b^* = a^*\tilde{b} + \tilde{a}b^* - \widetilde{ab}$. If we set $\bar{a} = a^* + \tilde{a}$ we have, therefore, $\bar{a}\bar{b} = a^*b^* + a^*\tilde{b} + \tilde{a}b^* = (ab)^* + \widetilde{ab} = \overline{ab}$, whence our result follows immediately.

3. Derivations.

Definition 3. 1. Let $\mathfrak{A}$ be any algebra (not necessarily associative); let $\mathfrak{B}$ be another algebra (of the same kind) which contains $\mathfrak{A}$ as a subalgebra. A linear mapping $a \rightarrow D(a)$ of $\mathfrak{A}$ into $\mathfrak{B}$ such that, for any elements $x, y \in \mathfrak{A}$, we have $D(x \cdot y) = D(x) \cdot y + x \cdot D(y)$ is called a derivation of $\mathfrak{A}$ into $\mathfrak{B}$. A derivation of $\mathfrak{A}$ into itself will sometimes be simply called a derivation of $\mathfrak{A}$.

If $\mathfrak{A}$ is an associative algebra contained in another associative algebra $\mathfrak{B}$, it is easy to verify that the mapping $a \rightarrow ba - ab$, where b is any fixed element of $\mathfrak{B}$, is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$. This derivation can trivially be extended to a derivation of $\mathfrak{B}$ into itself.

Similarly, if $\mathfrak{L}$ is a Lie algebra contained in another Lie algebra $\mathfrak{K}$, we see from Jacobi's identity that the mapping $l \rightarrow l \circ k$, where k is any fixed element of $\mathfrak{K}$, is a derivation of $\mathfrak{L}$ into $\mathfrak{K}$ which can be extended to a derivation of $\mathfrak{K}$ into itself.

Definition 3. 2. If $\mathfrak{A}$ is an associative algebra, the derivation $a \rightarrow a_0 a - a a_0$, where a_0 is any fixed element of $\mathfrak{A}$, is called an inner derivation of $\mathfrak{A}$.

If $\mathfrak{L}$ is a Lie algebra, the derivation $l \rightarrow l \circ l_0$, where l_0 is any fixed element of $\mathfrak{L}$, is called an inner derivation of $\mathfrak{L}$.

It is known that a semi-simple Lie algebra possesses only inner derivations.[3]

For associative algebras it has been shown[4] that, if $\mathfrak{A}$ is a semi-simple subalgebra of a normal simple algebra $\mathfrak{N}$, any derivation of $\mathfrak{A}$ into itself can be extended to an inner derivation of $\mathfrak{N}$.

Before we generalize these results we make the following definition:

Definition 3. 3. Let $\mathfrak{A}$ be an associative algebra, or a Lie algebra. $\mathfrak{A}$ is said to be reflexive if every derivation of $\mathfrak{A}$ into an arbitrary algebra $\mathfrak{B} \supseteq \mathfrak{A}$ can be extended to an inner derivation of $\mathfrak{B}$.

THEOREM 3. 1. *A Lie algebra is reflexive if and only if it possesses the property of Theorem* 2. 1.

Proof. 1). Suppose that a Lie algebra $\mathfrak{L}$ satisfies the condition of Theorem 2. 1. If $\mathfrak{K}$ is any Lie algebra containing $\mathfrak{L}$, we may regard it as a representation space of $\mathfrak{L}$, the transformation of $\mathfrak{K}$ corresponding to an element $l \in \mathfrak{L}$ being the mapping $k \rightarrow k \circ l$. If now D is any derivation of $\mathfrak{L}$ into $\mathfrak{K}$ we have $D(l \circ m) = D(l) \circ m + l \circ D(m) = D(l) \circ m - D(m) \circ l$, which shows that the mapping $l \rightarrow D(l)$ satisfies the conditions of Theorem 2. 1. Hence there exists an element $k_0 \in \mathfrak{K}$ such that $D(l) = k_0 \circ l$, whence we see that $\mathfrak{L}$ is reflexive. 2). Suppose that $\mathfrak{L}$ is reflexive. Let $\mathfrak{P}$ be a representation space of $\mathfrak{L}$. We define a Lie algebra $\mathfrak{K}$ whose underlying space is $\mathfrak{L} \times \mathfrak{P}$ by the commutation rule $(l, \vec{e}) \circ (m, \vec{f}) = (l \circ m,\ m \cdot \vec{e} - l \cdot \vec{f})$. (It is easy to verify that Jacobi's identity is satisfied). $\mathfrak{K}$ contains $(\mathfrak{L}, 0)$ as a subalgebra isomorphic with $\mathfrak{L}$.

---

[3] K. Yosida, "A characterization of the adjoint representations of the semi-simple Lie rings," *Japanese Journal of Mathematics*, vol. 14 (1938), p. 170.

[4] Cf. N. Jacobson, "Abstract derivation and Lie algebras," *Transactions of the American Mathematical Society*, vol. 42 (1937), p. 214.

Suppose now we have a mapping $l \to \vec{e}(l)$ of $\mathfrak{L}$ into $\mathfrak{P}$, as is considered in Theorem 2. 1. This gives a mapping $(l, 0) \to (0, \vec{e}(l))$ of $(\mathfrak{L}, 0)$ into $\mathfrak{K}$ which is a derivation since $(0, \vec{e}(l \circ m)) = (0, m \cdot \vec{e}(l) - l \cdot \vec{e}(m)) = (0, \vec{e}(l)) \circ (m, 0) + (l, 0) \circ (0, \vec{e}(m))$. But $(\mathfrak{L}, 0)$, being isomorphic with $\mathfrak{L}$, is reflexive, and hence there exists an element $(l_0, \vec{e}_0) \in \mathfrak{K}$ such that $(0, \vec{e}(l)) = (l, 0) \circ (l_0, \vec{e}_0)$, which implies that $\vec{e}(l) = l \cdot (-\vec{e}_0)$, as required in Theorem 2. 1. This completes the proof.

THEOREM 3. 2. *An associative algebra is reflexive if and only if it possesses the property of Theorem* 2. 2.

Proof. 1). If $\mathfrak{B}$ is any associative algebra containing $\mathfrak{A}$ we may regard $\mathfrak{B}$ as a two-sided $\mathfrak{A}$-module. A derivation D of $\mathfrak{A}$ into $\mathfrak{B}$ is a mapping of the type considered in Theorem 2. 2. The conclusion of Theorem 2. 2 becomes the statement that D is an inner derivation of $\mathfrak{L}$. Hence if $\mathfrak{A}$ has the property of Theorem 2. 2 then it is reflexive. 2). Suppose that $\mathfrak{A}$ is reflexive and let $\mathfrak{P}$ be a two-sided $\mathfrak{A}$-module. We construct an algebra $\mathfrak{B}$ over the linear space $\mathfrak{A} \times \mathfrak{P}$ by defining multiplication as follows: $(a, e) \cdot (b, f) = (ab, a \cdot \vec{f} + \vec{e} \cdot b)$. The rest of the proof is the same as the corresponding part of the proof of Theorem 3. 1.

From Theorems 3. 1, 2. 4, 2. 1, 1. 2 and 3. 2, 2. 3, 2. 2, we have now:

THEOREM 3. 3. *If an algebra (Lie or associative) is reflexive then every representation of it is completely decomposable.*

THEOREM 3. 4. *An algebra over a field of characteristic zero is semi-simple if and only if it is reflexive.*

Remarks. 1). We have shown without any assumption on the groundfield that if an algebra is reflexive then it is semi-simple. The converse was proved with the restriction that the groundfield be of characteristic zero. That this restriction is necessary for associative algebras is seen from results of N. Jacobson who considered the derivation algebras of certain fields of characteristic p. (Clearly, if a field were reflexive, its only derivation would be the mapping which maps every element on 0; but this is not the case for characteristic p). We refer to: N. Jacobson, "Abstract Derivation and Lie Algebras," *Transactions of the American Mathematical Society*, vol. 42 (1937), p. 217, for these results.

For Lie algebras, the necessity of our restriction is seen from the following example:

Let P be the field of the integers mod. 2 and define a Lie algebra of dimension 3 over P by the following equations of structure (x, y, z is a basis of the Lie algebra): $x \circ y = z$, $y \circ z = x$, $z \circ x = y$. The Lie algebra thus defined is clearly a simple Lie algebra. But the linear mapping D defined by the equations $D(x) = x$, $D(y) = y$, $D(z) = 0$ is a derivation which is not inner.

2). It may be useful to point out that the condition that an algebra $\mathfrak{A}$ be reflexive is genuinely stronger than the condition that all derivations of $\mathfrak{A}$ *into itself* be inner.

In fact, consider the Lie algebra $\mathfrak{L}$ of dimension 2 over (e. g.) the field of the rational numbers, with generators x, y, where $x \circ y = x$. It is easily verified that every derivation D of $\mathfrak{L}$ into itself is of the form $D(x) = \alpha x$, $D(y) = \beta x$. If we put $z = \alpha y - \beta x$ we have $D(x) = x \circ z$, $D(y) = y \circ z$, i. e. D is an inner derivation. But $\mathfrak{L}$ cannot be reflexive, since it is solvable.

Similarly, the associative algebra of all matrices of the form $\left(\begin{smallmatrix}\lambda & 0\\ \mu & \nu\end{smallmatrix}\right)$, with coefficients in the field of the rational numbers, is clearly not semi-simple and hence cannot be reflexive; but it is easily seen that every derivation D of this algebra into itself is of the form $D(E_{11}) = \alpha E_{21}$, $D(E_{31}) = \beta E_{21}$, $D(E_{22}) = -\alpha E_{21}$, where we have written E_{11} for the matrix $\left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right)$, etc. Now, if $a_0 = \alpha E_{21} + \beta E_{22}$, we $D(a) = a_0 a - a a_0$ for every element a of this algebra, i. e. D is an inner derivation.

4. Derivation algebras. If $\mathfrak{A}$ is any algebra, the derivations $D_1, D_2, \cdots$ of $\mathfrak{A}$ constitute a Lie algebra of linear transformations, commutation being defined by the formula $D_1 \circ D_2 = D_2 D_1 - D_1 D_2$. The Lie algebra thus obtained is called the derivation algebra of $\mathfrak{A}$. We propose to study the structure of derivation algebras.

Definition 4.1. A subset $\mathfrak{J}$ of an algebra (Lie or associative) is called characteristic if $D(\mathfrak{J}) \subseteq \mathfrak{J}$ for every derivation D of the algebra.

For example, it is easy to verify that the center of a Lie algebra is always a characteristic ideal. Similarly, the center of an associative algebra is always a characteristic subalgebra.

We shall require the following:

THEOREM 4.1.[5] *Let $\mathfrak{L}$ be a Lie algebra over a field of characteristic* 0; *let $\mathfrak{S}$ be its radical (the maximal solvable ideal). Then $\mathfrak{S}$ is a characteristic ideal.*

---

[5] This was first proved by H. Zassenhaus, *Hamb. Abhandlungen*, Band. 13 (1940), p. 79.

Proof. Let $\mathfrak{H} \supset \mathfrak{H}^{(1)} \supset \mathfrak{H}^{(2)} \supset \cdots \mathfrak{H}^{(p+1)} = \{0\}$ be the sequence of the successive derived algebras of $\mathfrak{H}$. Let D be any derivation of $\mathfrak{L}$. Suppose that $D^i(\mathfrak{H}^{(k+1)}) \subseteq \mathfrak{H}$ for all $i = 1, 2, \cdots$; (this holds trivially for $k = p$). Then we shall show that $D^i(\mathfrak{H}^{(k)}) \subseteq \mathfrak{H}$; $i = 1, 2, \cdots$.

Since $\mathfrak{H}^{(k)}$ is an ideal of $\mathfrak{L}$ it is easy to see that the set $[\mathfrak{H} + D(\mathfrak{H}^{(k)})]$ of all elements of the form $h + D(h')$, where $h \in \mathfrak{H}$ and $h' \in \mathfrak{H}^{(k)}$ constitutes an ideal in $\mathfrak{L}$. If $h_1, h_2 \in \mathfrak{H}^{(k)}$ we have $2\,D(h_1) \circ D(h_2) \equiv D^2(h_1 \circ h_2)$ mod. $\mathfrak{H}$, and it follows from our inductive hypothesis that $D(h_1) \circ D(h_2) \in \mathfrak{H}$.

Hence the derived algebra $[\mathfrak{H} + D(\mathfrak{H}^{(k)})]'$ is contained in $\mathfrak{H}$: $[\mathfrak{H} + D(\mathfrak{H}^{(k)}]$ is therefore a solvable ideal of $\mathfrak{L}$, and since $\mathfrak{H}$ is maximal we must have $D(\mathfrak{H}^{(k)}) \subseteq \mathfrak{H}$.

Suppose we have already proved that $D^i(\mathfrak{H}^{(k)}) \subseteq \mathfrak{H}$ for all $i < n$. Then $[\mathfrak{H} + D^n(\mathfrak{H}^{(k)})]$ is an ideal in $\mathfrak{L}$, for $D^n(h) \circ x \equiv D^n(h \circ x)$ mod. $\mathfrak{H}$ for all $h \in \mathfrak{H}^{(k)}$ and $x \in \mathfrak{L}$. The derived algebra $[\mathfrak{H} + D^n(\mathfrak{H}^{(k)})]'$ consists of sums of elements of $\mathfrak{H}$ and elements of the form $D^n(h_1) \circ D^n(h_2)$, with $h_1, h_2 \in \mathfrak{H}^{(k)}$. But $D^n(h_1) \circ D^n(h_2) \equiv 1/{}_{2n}C_n \cdot D^{2n}(h_1 \circ h_2)$ mod. $\mathfrak{H}$, whence we may conclude, as before, that we must have $D^n(\mathfrak{H}^{(k)}) \subseteq \mathfrak{H}$. Thus we have shown that, if $D^i(\mathfrak{H}^{(k+1)}) \subseteq \mathfrak{H}$ for all i, then also $D^i(\mathfrak{H}^{(k)}) \subseteq \mathfrak{H}$ for all i, and our theorem follows by induction on k.

In exactly the same manner we can prove:

THEOREM 4. 2. *The radical of an associative algebra over a field of characteristic 0 is a characteristic ideal.*

Let now $\mathfrak{A}$ be a Lie algebra or an associative algebra over a field of characteristic 0. In either case we may write $\mathfrak{A} = \mathfrak{H} + \mathfrak{K}$, where $\mathfrak{H}$ is the radical of $\mathfrak{A}$ and $\mathfrak{K}$ is a semi-simple subalgebra. If D is any derivation of $\mathfrak{A}$ we have $D(\mathfrak{H}) \subseteq \mathfrak{H}$. On the other hand we know that there exists an element $a_0 \in \mathfrak{A}$ such that $D(k) = k \circ a_0$ for every $k \in \mathfrak{K}$. (In the associative case, $k \circ a_0$ will stand for the expression $a_0 k - k a_0$). Let D_{a_0} denote the inner derivation effected by a_0. If we put $D' = D - D_{a_0}$ we have $D'(\mathfrak{K}) = \{0\}$. We have, then, $D'(h \circ k) = D'(h) \circ k$ (for Lie algebras) or

$$\left.\begin{cases} D'(kh) = kD'(h) \\ D'(hk) = D'(h)k \end{cases}\right\} \text{(for associative algebras).} \qquad \ldots\ldots\ldots\ldots (A)$$

Conversely, any derivation of $\mathfrak{H}$ satisfying the condition (A) gives a derivation of $\mathfrak{A}$ if we define $D(k) = 0$ for every $k \in \mathfrak{K}$. Hence we have:

THEOREM 4. 3. *Every derivation of $\mathfrak{A}$ is the sum of an inner derivation and a derivation which annuls $\mathfrak{K}$.*

A derivation of $\mathfrak{H}$ can be extended to a derivation of $\mathfrak{A}$ if and only if it is the sum of an inner derivation (by means of an element of $\mathfrak{A}$, not necessarily of $\mathfrak{S}$) and a derivation with the property (A).

We are now in a position to prove the following result:

THEOREM 4. 4. *The derivation algebra $\mathfrak{D}$ of a Lie algebra $\mathfrak{L}$ over a field of characteristic* 0 *is semi-simple if and only if $\mathfrak{L}$ is semi-simple.*

Proof. If $\mathfrak{L}$ is semi-simple, $\mathfrak{D}$ consists only of inner derivations. Moreover, since the center of $\mathfrak{L}$ reduces to $\{0\}$, $\mathfrak{D}$ is isomorphic with $\mathfrak{L}$ and hence is semi-simple.

Suppose now that $\mathfrak{D}$ is semi-simple. Let us write $\mathfrak{L} = \mathfrak{S} + \mathfrak{K}$, as before. Denote by $\mathfrak{D}_{\mathfrak{S}}$ the set of the inner derivations of $\mathfrak{L}$ which are effected by the elements of $\mathfrak{S}$. If D_k is such a derivation, and D is any other derivation of $\mathfrak{L}$ we have $D_k \circ D = D_{D(k)}$, whence it is clear that $\mathfrak{D}_{\mathfrak{S}}$ is an ideal in $\mathfrak{D}$. Since $\mathfrak{S}$ is solvable, it follows moreover that $\mathfrak{D}_{\mathfrak{S}}$ is solvable, and hence must reduce to $\{0\}$. This means that $\mathfrak{S}$ is the center of $\mathfrak{L}$. Since every derivation is inner as far as $\mathfrak{K}$ is concerned, this implies that every derivation maps $\mathfrak{K}$ into itself. Finally, every derivation of the center $\mathfrak{S}$ can be extended to a derivation of $\mathfrak{L}$ which annuls $\mathfrak{K}$, and the same statement is true if $\mathfrak{S}$ and $\mathfrak{K}$ are interchanged. Every derivation is the sum of derivations obtained in this manner; in fact, $\mathfrak{D}$ is isomorphic with the direct sum of the derivation algebra of $\mathfrak{S}$ and the derivation algebra of $\mathfrak{K}$. Since $\mathfrak{S}$ is abelian, its derivation algebra is the Lie algebra of all linear transformations of $\mathfrak{S}$, and this is clearly not semi-simple unless $\mathfrak{S} = \{0\}$. This completes the proof.

With regard to associative algebras it is known that the derivation algebra of a normal simple algebra is simple (N. Jacobson, "Abstract Derivation and Lie Algebras," *Transactions of the American Mathematical Society,* vol. 42). We shall prove:

THEOREM 4. 5. *Let $\mathfrak{A}$ be an associative algebra over a field of characteristic* 0, *and let $\mathfrak{D}$ be the derivation algebra of $\mathfrak{A}$. Then $\mathfrak{D}$ is semi-simple or $\{0\}$ if and only if $\mathfrak{A}$ is semi-simple.*

Proof. If $\mathfrak{A}$ is semi-simple its derivation algebra consists only of inner derivations. Let us denote by $\mathfrak{A}_l$ the Lie algebra obtained from $\mathfrak{A}$ by defining the commutator of two elements as $a \circ b = ba - ab$. Then it is clear that the derivation algebra of $\mathfrak{A}$ is isomorphic with $\mathfrak{A}_l/\mathfrak{Z}$, where $\mathfrak{Z}$ is the center of $\mathfrak{A}_l$. Now, it is known that the derived algebra $\mathfrak{A}'_l$ of $\mathfrak{A}_l$ is semi-simple or $\{0\}$.[6]

---

[6] W. Landherr, "Über einfache Lie-sche Ringe," *Hamb. Abhandlungen,* Band 11 (1935), pp. 41-64. N. Jacobson, "Simple Lie algebras over a field of characteristic 0," *Duke Mathematical Journal,* vol. 4 (1938), p. 536.

Let $\mathfrak{H}$ be the radical of $\mathfrak{A}_l$. Then $\mathfrak{H} \circ \mathfrak{A}_l$ is a solvable ideal in $\mathfrak{A}'_l$ and hence $\mathfrak{H} \circ \mathfrak{A}_l = \{0\}$. But this means that $\mathfrak{H} \subseteq \mathfrak{Z}$ (in fact $\mathfrak{H} = \mathfrak{Z}$). Hence $\mathfrak{A}_l/\mathfrak{Z}$ is semi-simple or $\{0\}$, which proves the first part of our theorem.

Suppose now that $\mathfrak{D}$ is semi-simple or $\{0\}$. Write $\mathfrak{A} = \mathfrak{N} + \mathfrak{K}$, where $\mathfrak{K}$ is a semi-simple subalgebra and $\mathfrak{N}$ the radical of $\mathfrak{A}$. It follows, then, as in the proof of Theorem 4. 4, that $\mathfrak{N}$ is contained in the center of $\mathfrak{A}$ and hence that $\mathfrak{K}$ is mapped into itself by every derivation of $\mathfrak{A}$. Let $\mathfrak{N} \supset \mathfrak{N}^2 \supset \cdots \supset \mathfrak{N}^k = \{0\}$ be the series of the successive powers of $\mathfrak{N}$. It is easily seen by an induction on the exponent i that every $\mathfrak{N}^i$ is a characteristic ideal of $\mathfrak{A}$. Suppose that $\mathfrak{N} \neq \{0\}$.

If $\mathfrak{N}^2 = \{0\}$ we can define a derivation D of $\mathfrak{A}$ as follows:

$$D(r) = r \quad \text{if } r \in \mathfrak{N}; \qquad D(k) = 0 \quad \text{if } k \in \mathfrak{K}.$$

It is clear that D is really a derivation of $\mathfrak{A}$. If D^* is any other derivation of $\mathfrak{A}$ we have

$$(D \circ D^*)(r) = (D^*D - DD^*)(r) = D^*(r) - D(D^*(r)) = 0 \text{ if } r \in \mathfrak{N}$$
$$(\text{since } D^*(r) \in \mathfrak{N})$$

and $$(D \circ D^*)(k) = \dots\dots\dots\dots\dots = -D(D^*(k)) = 0 \text{ if } k \in \mathfrak{K}$$
$$(\text{since } D^*(k) \in \mathfrak{K}).$$

Hence $D \circ D^* = 0$ for every $D^* \in \mathfrak{D}$, which contradicts the hypothesis that $\mathfrak{D}$ is semi-simple or $\{0\}$. Hence $\mathfrak{N}^2 \neq \{0\}$, and so, in the series above, $k > 2$. If $r_0 \in \mathfrak{N}^{k-2}$ we can define a derivation D_{r_0} of $\mathfrak{A}$ as follows:

$$D_{r_0}(r) = r_0 \cdot r \quad \text{if } r \in \mathfrak{N}; \qquad D_r(k) = 0 \quad \text{if } k \in \mathfrak{K}.$$

(Again, the verification that this is really a derivation is trivial.)

If D is any other derivation of $\mathfrak{A}$ we have

$$(D_{r_0} \circ D)(r) = D(r_0 r) - r_0 \cdot D(r) = D(r_0) \cdot r \qquad \text{if } r \in \mathfrak{N},$$
$$(\text{since } D(r) \in \mathfrak{N})$$

and $$(D_{r_0} \circ D)(k) = -D_{r_0}(D(k)) = 0 \qquad \text{if } k \in \mathfrak{K}, \ (\text{since } D(k) \in \mathfrak{K}).$$

Hence $D_{r_0} \circ D = D_{D(r_0)}$, which shows that the derivations of the form D_{r_0}, $r_0 \in \mathfrak{N}^{k-2}$, constitute an ideal in $\mathfrak{D}$. Since there is at least one non-zero derivation of this type (otherwise we should have $\mathfrak{N}^{k-1} = \{0\}$), and since the ideal just defined is clearly abelian, we have again reached a contradiction. Hence $\mathfrak{N} = \{0\}$ and $\mathfrak{A}$ is semi-simple.

PRINCETON UNIVERSITY.

THE LIMIT OF THE RATIO OF ARC TO CHORD.*

By George Comenetz.

1. Introduction. This paper deals with a problem in the geometry of Euclidean space of three complex dimensions. Given a fixed point O on an analytic curve in this space, let V be a variable point which approaches O along the curve. The problem is to find the limit of the ratio of the arc OV to the chord OV. If the space were real the limit would of course always be unity,[1] and even in complex space the limit is unity whenever the direction of the curve at O is not isotropic, that is, whenever the tangent line at O is not an isotropic line. But if the direction is isotropic the limit as a rule is a complex number other than unity, as Kasner discovered.[2] For the corresponding problem in two dimensions he found a theorem, quoted below as Theorem I, which shows how the limit is related to standard geometrical quantities connected with the curve. In Theorem II we present a similar result for curves in three dimensions.

After the flat spaces of two and three dimensions, the question is considered in a curved two-space, namely a surface lying in the given Euclidean three-space. A curve is drawn on the surface, passing through a given point O. What is there about the geometry of the surface at the point, and of the curve, which determines the value of the limit? Then, what kind of a set of limits can be obtained at a given point on a surface by using all possible curves? The latter question was answered by Kasner.[3] An answer to the former, based on Theorem II, is given below in Sections **3** and **7**.

If the curve under investigation is not isotropic at the given point O, the limit is unity as already mentioned. If the curve is isotropic at O and also throughout a neighborhood of O, that is, if it is an isotropic curve, the limit is zero, for the arc vanishes identically (and the chord does not, for we assume the curve is not an isotropic straight line; if it were, both arc and chord would vanish and no ratio could be formed). The determination of the limit

---

* Received August 29, 1941. Abstracts in *Bulletin of the American Mathematical Society*, vol. 38 (1932), p. 192; vol. 45 (1939), p. 87.

[1] This theorem dates back at least to Newton's *Principia*, Book I, Lemma VII.

[2] "The ratio of the arc to the chord of an analytic curve need not approach unity," *Bulletin of the American Mathematical Society*, vol. 20 (1914), pp. 524-531.

[3] "Complex Geometry and Relativity: Theory of the 'Rac' Curvature," *Proceedings of the National Academy of Sciences*, vol. 18 (1932), pp. 267-274.

therefore becomes a problem only when the curve is isotropic at O itself, but at no other point within a neighborhood of O. This is interesting because the curves studied in complex differential geometry are usually either not at all or else wholly isotropic, whereas here the attention centers exactly on the isolated isotropic points of ordinary, non-isotropic curves.

Somewhat analogous to this is the fact that in the discussion of curves on a surface two cases must be treated, depending on whether the tangent plane to the surface at O is an isotropic plane or not; and the case in which it is isotropic is considerably the more complicated of the two.

In Section **6** we investigate the isotropic net of curves on a surface, at a point where the tangent plane is isotropic. Such a point is singular for the differential equation $ds^2 = 0$ of the isotropic net. The various things that can happen are described in Theorems III-V.

The theorem of Kasner on the geometrical determination of the limit for curves in the plane is as follows.

THEOREM I. *Let O be a regular point on an analytic curve in space of two complex dimensions. If the direction of the curve at O is not isotropic the limit of the ratio of arc to chord is* 1. *If the direction at O is isotropic, let n be the order of contact of the curve with its isotropic tangent line at O. Then the limit of the ratio is* $2(n+1)^{\frac{1}{2}}/(n+2)$.

The corresponding theorem for three dimensions is this:

THEOREM II. *Let O be a regular point on an analytic curve in space of three complex dimensions. If the direction of the curve at O is not isotropic the limit of the ratio of arc to chord is* 1. *If the direction at O is isotropic, let n be the order of contact of the curve with its isotropic tangent line at O, and let m be the order of contact of the curve with the isotropic plane which contains that isotropic line. Then if $m > 2n$ the limit of the ratio is* ± 1, *and if $m < 2n$ the limit is* $2(m+1)^{\frac{1}{2}}/(m+2)$. *If $m = 2n$, isotropic curves exist which have contact of order m with the given curve at O. If there exists in addition an isotropic curve having contact of order higher than m with the given curve, the limit is zero. If no such isotropic curve of higher contact exists, the limit equals*

$$\delta/[\delta^2 - (m+2)^2/4(m+1)]^{\frac{1}{2}},$$

where δ is the value at O of the derivative of the radius of curvature with respect to arc-length. For a suitable curve this can become any complex number other than ± 1, $2(m+1)^{\frac{1}{2}}/(m+2)$ *and zero.*

Thus in three dimensions, as in two, the limit depends on orders of contact, which are integer invariants; but in three dimensions it may also depend on a differential invariant, the derivative of the radius of curvature with respect to arc-length.

2. Proof of Theorem II. We assume as known that the limit of the ratio of arc to chord is 1 when the curve's direction is not isotropic.[4]

Suppose, therefore, that the direction of the curve at O is isotropic. We may take O as the origin and the isotropic straight line $y=0$, $z=ix$ as the tangent line at O. Then the equations of the curve in some neighborhood of O may be expressed by two power series,

$$y = a_{\nu+1}x^{\nu+1} + a_{\nu+2}x^{\nu+2} + \cdots, \tag{1}$$

$$z = ix + b_{m+1}x^{m+1} + b_{m+2}x^{m+2} + \cdots, \tag{2}$$

where $a_{\nu+1} \neq 0$, $b_{m+1} \neq 0$, $\nu \geqq 1$ and $m \geqq 1$ (if $y \equiv 0$ we say that $\nu = \infty$, and then there is no $a_{\nu+1}$; similarly if $z \equiv ix$, $m = \infty$ and there is no b_{m+1}).

The isotropic plane which contains the isotropic line $y=0$, $z=ix$ is the plane $z=ix$. The integer m represents the order of contact of the curve with this plane at O. The order of contact n of the curve with its isotropic tangent line is the smaller of the two integers m and ν or either one if they are equal, that is,

$$n = \min(m, \nu).$$

First suppose that $m > 2n$, in which case n is ν, so that $m > 2\nu$. Let c stand for the chord and s for the arc, each measured from O to a variable point (x, y, z) on the curve. Then

$$c = (x^2 + y^2 + z^2)^{\frac{1}{2}}, \qquad s = \int_0^x (1 + y'^2 + z'^2)^{\frac{1}{2}} dx.$$

By substitution from (1) and (2),

$$\begin{aligned}
c &= [x^2 + (a^2_{\nu+1}x^{2\nu+2} + \cdots) + (-x^2 + 2ib_{m+1}x^{m+2} + \cdots)]^{\frac{1}{2}},\\
c &= (a^2_{\nu+1}x^{2\nu+2} + \cdots)^{\frac{1}{2}} = \pm(a_{\nu+1}x^{\nu+1} + \cdots);\\
s &= \int_0^x \{1 + [(\nu+1)^2 a^2_{\nu+1}x^{2\nu} + \cdots]\\
&\qquad\qquad + [-1 + 2i(m+1)b_{m+1}x^m + \cdots]\}^{\frac{1}{2}} dx,\\
s &= \pm \int_0^x [(\nu+1)a_{\nu+1}x^\nu + \cdots]dx = \pm\ (a_{\nu+1}x^{\nu+1} + \cdots).
\end{aligned}$$

---

[4] In this case the arc and chord are functions of the parameter of the curve each of which separates into two distinct branches in a neighborhood of the given point. It is understood that corresponding branches of these two functions are selected to form the ratio of arc to chord, corresponding branches being those which have the same derivative at the given point. If such a selection is not made the ratio can approach both $+1$ and -1. However in the remaining discussion, where the curve's direction is isotropic,

Dividing s by c (the two $\pm$ signs are independent), and writing R for the limit of s/c as x approaches zero, we see that

$$R = \pm 1$$

as Theorem II states.

The next hypothesis is that $m < 2n$. Then $m < 2\nu$. Calculating as above we have

$$c = (2ib_{m+1})^{\frac{1}{2}}x^{(m+2)/2} + \cdots,$$

$$s = \int_0^x \{[2i(m+1)b_{m+1}]^{\frac{1}{2}}x^{m/2} + \cdots\}\,dx =$$

$$= \frac{2}{m+2}[2i(m+1)b_{m+1}]^{\frac{1}{2}}x^{(m+2)/2} + \cdots,$$

$$R = 2\sqrt{m+1}/(m+2).$$

This confirms the third result in Theorem II.

Suppose finally that $m = 2n$, which means that $m = 2\nu$. In this case

$$c = (a^2{}_{\nu+1} + 2ib_{m+1})x^{(m+2)/2} + \cdots,$$

$$s = \frac{2}{m+2}[(\nu+1)^2a^2{}_{\nu+1} + 2i(m+1)b_{m+1}]^{\frac{1}{2}}x^{(m+2)/2} + \cdots. \tag{3}$$

The leading coefficients cannot both vanish as that would require m to be zero. Hence

$$R = \frac{2}{m+2}\left[\frac{(\nu+1)^2a^2{}_{\nu+1} + 2i(m+1)b_{m+1}}{a^2{}_{\nu+1} + 2ib_{m+1}}\right]^{\frac{1}{2}}. \tag{4}$$

(If $a^2{}_{\nu+1} + 2ib_{m+1} = 0$, then $R = \infty$.)

The fact that there exist isotropic curves $Y = Y(x)$, $Z = Z(x)$ which have contact of order m with the given curve may be recognized as follows. For $Y(x)$ any convergent power series in x may be chosen which agrees with (1) up to and including the term $a_{2\nu}x^{2\nu}$. Then $Z'(x)$ is determined by the condition for an isotropic curve,

$$1 + Y'^2 + Z'^2 = 0 \qquad \text{or} \qquad Z' = (-1 - Y'^2)^{\frac{1}{2}}.$$

From this we compute that

$$Z(x) = ix + \frac{i(\nu+1)^2}{2(2\nu+1)}\,a^2{}_{\nu+1}x^{2\nu+1} + \cdots. \tag{5}$$

As this series agrees with (2) through the first 2ν or m terms, the isotropic curve $Y = Y(x)$, $Z = Z(x)$ touches the given curve (1), (2) to the order m.

---

we shall not try to avoid ambiguity of sign. The limit in every case will then be indeterminate as to sign. Thus all the square root symbols are to be understood ambiguously. If we were to consider the square of the ratio of arc to chord instead of the ratio itself, then there would always be a unique limit.

On the other hand, an isotropic curve having contact of order higher than m with the given curve will exist if, and only if, the coefficient of $x^{2\nu+1}$ in (5) equals the coefficient of x^{m+1} in (2). This condition may be written as

$$H \equiv (\nu+1)^2 a^2{}_{\nu+1} + 2i(m+1)b_{m+1} = 0.$$

Now returning to (4), we see that if an isotropic curve exists which has contact of order higher than m with the given curve, so that $H = 0$, then R will be zero.

Suppose, on the contrary, that there is no isotropic curve which touches the given curve to an order higher than m. Hence H is not zero. From the ordinary formula

$$r^2 = \frac{(1+y'^2+z'^2)^3}{(1+z'^2)y''^2+(1+y'^2)z''^2-2y'z'y''z''}$$

for the radius of curvature r, we can then calculate that

$$r^2 = \frac{iH^2}{2b_{m+1}\nu^2(m+1)}x^{m+2}+\cdots \tag{6}$$

for the given curve (1), (2).[5] The limit of r/s as x approaches zero is δ, the derivative of the radius of curvature with respect to arc-length at O, because r and s both vanish at O (similarly the limit of the ratio of arc to chord is also the derivative of the arc with respect to the chord). From (3) and (6) therefore,

$$\delta = \frac{m+2}{m}\left[\frac{iH}{2(m+1)b_{m+1}}\right]^{\frac{1}{2}}. \tag{7}$$

This can be solved for the ratio $a^2{}_{\nu+1}/b_{m+1}$. Substituting the result in (4) and simplifying, we find that

$$R = \frac{\delta}{\left[\delta^2-\dfrac{(m+2)^2}{4(m+1)}\right]^{\frac{1}{2}}}, \tag{8}$$

which is the formula given in Theorem II.

By assumption $a_{\nu+1}$, b_{m+1} and H do not vanish, but $a_{\nu+1}$ and b_{m+1} are otherwise unrestricted. It follows that δ in (7) may have any finite complex value with the exception of zero and $\pm i(m+2)/m$. Now from (8), R^2 is a linear fractional expression in δ^2. Hence R can be any complex number, including ∞, with the exception of zero, ± 1 and $2\sqrt{m+1}/(m+2)$. This ends the proof of Theorem II.

---

[5] The radius of curvature in the sense of limit of the ratio of arc to angle-difference does not exist at O. We define r to be zero at O, zero being the limit of r as a point on the curve approaches O.

The theorem implies that if the limit of the ratio of arc to chord is known to be zero for a certain curve at a certain point, then the order of contact of the curve with its (necessarily isotropic) tangent line at the point is half the order of contact with the isotropic plane containing the tangent line. In particular this is true at every point of an isotropic curve. But it is easy to show that directly, by applying the condition $1 + y'^2 + z'^2 = 0$ to (1), (2).

3. Surface with non-isotropic tangent plane. Let O be a given regular point on a given analytic surface S. Let C be an arbitrary analytic curve on S, regular at O.

The limit R of the ratio of arc to chord of C at O is determined by the curve's m and n or its δ or its behavior relative to the neighboring isotropic curves, all as described in Theorem II. But C is now not a general curve through O, but a member of the restricted class of curves that lie on S. Its m and n and so on are therefore, in turn, at least partly determined by S. To describe this determination and the consequent determination of R in geometrical terms is the main object of the rest of the paper. The figures and quantities we employ may be of use in other problems in the geometry of surfaces in which orders of contact are significant.

Let T be the tangent plane to S at O. In this section we treat the case in which T is non-isotropic. In Section **4** we define an invariant analogous to order of contact, preliminary to a discussion in the remaining sections of the rather involved geometry when T is isotropic.

As T is now assumed non-isotropic, there lie in it two isotropic straight lines passing through O. Let these be called L and L' and let their respective orders of contact with S be p and p'.

If C is tangent neither to L nor to L', the limit R is 1.

Suppose now that C is tangent to L. Draw the plane M which passes through L and through the normal to S at O. This is the isotropic plane that contains L. Let the curve of intersection of M with S be named D. Since the order of contact of L with S is p, the order of contact of L with D is also p.

Let m denote the order of contact of C with D. The order of contact of C with plane M is the same m (this can be deduced from the definitions of the order of contact of a curve with a curve and a curve with a surface, and the fact that a line through a point of C and a point of D approaches parallelism to T as the two points approach O). Let n be the order of contact of C and L.

If α, β, γ are three curves tangent at a point, and if $\alpha\beta$, $\beta\gamma$, $\alpha\gamma$ denote their orders of contact one with another, there is the obvious triangle inequality

$$\alpha\gamma \geqq \min(\alpha\beta, \beta\gamma), \tag{9}$$

in which the equality sign holds if $\alpha\beta \neq \beta\gamma$ and may or may not hold if $\alpha\beta = \beta\gamma$.

Applying this to C, D and L we have

$$(10) \qquad n = \min(m, p) \text{ if } m \neq p, \qquad n \geqq m \text{ or } p \text{ if } m = p.$$

There are three cases, depending on the size of m as compared with $2p$.

First, if $m < 2p$, it follows from (10) that $m < 2n$; for if $m \leqq p$, then $n \geqq m$ and hence $2n > m$; if $p < m < 2p$, then $n = p$ and again $2n > m$. Second, if $m = 2p$, we see that $n = p$ and therefore $m = 2n$. Third, if $m > 2p$, then $n = p$ and $m > 2n$. Thus when m is less than, equal to or greater than $2p$, m is correspondingly less than, equal to or greater than $2n$.

Let us use the abbreviation

$$R_h \equiv 2[h+1]^{\frac{1}{2}}/(h+2).$$

By Theorem II, when $m < 2p$, R has the value R_m. When $m > 2p$, R is ± 1. Suppose finally that $m = 2p$.

The surface S contains just one isotropic curve, which we call E, that passes through O in the direction of L. It is that member of the isotropic net of curves $ds^2 = 0$ on S which is determined by the direction element O, L. The order of contact of E with M must be double its order of contact with L, by the remark at the end of Section 2. As we saw just above, a curve on the surface can have this property only if its order of contact with D is $2p$. Consequently E has contact of order $2p$ with D.

If an isotropic curve in the L direction has contact of order higher than $2p$ with S, it has contact of order higher than $2p$ with E (it is even true that if it has contact of order higher than p with S, it has that same order of contact with E). We omit the short calculation which proves this.

If C, with m equal to $2p$, is drawn so as to coincide with E or at least to have contact of order higher than $2p$ with E, R will be zero by Theorem II. If the order of contact of C with E is exactly $2p$, the minimum possible value in virtue of the triangle inequality applied to C, D and E, then C cannot have contact of order higher than $2p$ with any isotropic curve. If it did, the latter would have contact of order higher than $2p$ with S, but exactly $2p$ with E on account of the triangle inequality, in contradiction with the previous paragraph. By Theorem II, R then equals

$$\delta/[\delta^2 - 1/R^2_{2p}]^{\frac{1}{2}},$$

an expression which never has the values ± 1, R_{3p}, or zero. For a suitable curve C it can have any other complex value, but because C is confined to S

this does not follow from Theorem II. It requires a separate proof, which we omit, showing that δ can still have any finite value other than zero and $\pm i(p+1)/p$.

Summary. There are the surface S, the point O, and the isotropic line L through O having contact of order p with S. There is the curve D on S, the normal projection of L upon S. There is the isotropic curve E on S, having contact of order $2p$ with D. These are all the fixed elements.

Draw a curve C on S touching D to the first order at O. Now imagine C, always remaining on the surface, to be altered gradually so that its order of contact CD or m with D rises steadily through the integers $1, 2, 3, \cdots$ until finally C merges with D. The progress of C will be accompanied by changes in the value of the limit R according to the following rule:

$$\begin{aligned} CD \text{ or } m < 2p &: R = R_m; \\ CD = 2p,\ CE = 2p &: R = \delta/[\delta^2 - 1/R^2{}_{2p}]^{\frac{1}{2}}, \text{ which may be any number} \\ &\qquad\qquad\qquad\qquad\qquad \text{but } \pm 1,\ R_{2p} \text{ and } 0; \\ CD = 2p,\ CE > 2p &: R = 0; \\ CD > 2p &: R = \pm 1. \end{aligned}$$

Of course if $p = \infty$, that is if L lies in S, only the first stage occurs.

Curves touching the other line L' behave in the same way, p being replaced by p' and D and E by correspondingly constructed curves D' and E'.

4. Order of coincidence of directions.

In order to investigate the case of a surface with an isotropic tangent plane, we find it desirable to define an arithmetic invariant of a family of directions, analogous to the order of contact of a curve with its tangent line.

Let there be given a one-parameter family of directions $l(t)$, which we may represent as straight lines through the origin O, having direction numbers which are analytic in the parameter t. Suppose that as t approaches a certain value t_0 the lines approach a certain line $l(t_0)$, and furthermore suppose that the plane determined by $l(t)$ and $l(t_0)$ approaches a certain plane J. Take a fixed point P on $l(t_0)$ distinct from O and pass a fixed plane K through P not containing $l(t_0)$. Let $l(t)$ cut K at a point $Q(t)$. In K draw a fixed non-isotropic line j through P. Draw a line from $Q(t)$ to a point $N(t)$ on j, in some fixed direction not parallel to J. If μ is the order of the infinitesimal distance PN relative to $t - t_0$, we shall say that the *order of coincidence* of the directions $l(t)$ with the limiting direction $l(t_0)$ is μ. (If $l(t) \equiv l(t_0)$, $\mu = \infty$.)

Suppose that $e(t), f(t), g(t)$ are direction numbers of $l(t)$, approaching the direction numbers $e(t_0), f(t_0), g(t_0)$ of $l(t_0)$ as t approaches t_0. Then μ

equals the order in $t-t_0$ of that one (or those) of the three two-rowed determinants in the matrix

$$(11)\qquad \left\|\begin{matrix} e(t_0) & f(t_0) & g(t_0) \\ e(t) & f(t) & g(t) \end{matrix}\right\|$$

which has (or have) lowest order in $t-t_0$. Thus μ does not depend on the choice of P, K, j and the direction of QN. We omit the proof of this formula.

The integer μ may not be purely geometric, since it depends on the parameter t. But if t has geometric meaning then μ does become purely geometric. For instance, if the $l(t)$ are a set of directions associated with the points of a curve (they then form a "series"), and if t is arc-length along the curve, μ is geometric. Even if t is merely the x-coördinate employed as parameter along the curve, μ is still geometric, the reason being that for any x-coördinate which can be used as parameter $x-x_0$ is of the first order relative to the $x-x_0$ for any other, so that the order of PN is independent of the particular x-coördinate. For example, if the $l(t)$ are the tangent directions along the curve and t is the x-coördinate, μ equals the order of contact of the curve with its tangent line at t_0, as is easily proved.

If neither $l(t_0)$ nor J is isotropic, μ is simply the order of the infinitesimal angle between $l(t_0)$ and $l(t)$ relative to $t-t_0$.

5. Surface with isotropic tangent plane. Once more let O be a given point on a given surface S and let C be any curve on S passing through O. We now suppose that the plane T which is tangent to S at O is an isotropic plane; it therefore contains a single isotropic line L passing through O. Let p be the order of contact of L with S at O. Let F stand for an arbitrary one of the curves on S passing through O and having contact of maximum order, that is of order p, with L.

We may take O as the origin and assume $z=ix$ as the equation of T. Then L is the line $z=ix$, $y=0$. The equation of the surface, arranged according to powers of y, may be written as

$$(12)\qquad \begin{aligned} z = (ix + c_{p+1}x^{p+1} + c_{p+2}x^{p+2} + \cdots) \\ + (d_\lambda x^\lambda + d_{\lambda+1}x^{\lambda+1} + \cdots)y + (\cdots)y^2 + \cdots \end{aligned}$$

where $p \geqq 1$, $\lambda \geqq 1$, $c_{p+1} \neq 0$ and $d_\lambda \neq 0$ (if there are no terms after ix in the first parentheses there is no c_{p+1}, L lies in S, and we say that $p=\infty$; similarly it may be that $\lambda=\infty$).

The equations of a curve on S tangent to L are

$$(13)\qquad y = a_{\nu+1}x^{\nu+1} + a_{\nu+2}x^{\nu+2} + \cdots,$$

$$(14)\qquad z = (ix + c_{p+1}x^{p+1} + \cdots) + (a_{\nu+1}d_\lambda x^{\nu+\lambda+1} + \cdots) + \cdots,$$

where $\nu \geqq 1$ and $a_{\nu+1} \neq 0$ (or ν may be ∞); (14) is obtained by substituting (13) into (12). No power of x lower than $x^{2\nu+2}$ appears among the terms symbolized by the last three dots.

The order of contact n of this curve with L is given by

$$n = \min(\nu, p). \tag{15}$$

The curves F are therefore represented by (13) and (14) when $\nu \geqq p$. It follows that they have contact of order p with the tangent plane T.

Any two of the F-curves have contact of order at least p with one another, and, conversely, any curve on S which has contact of order at least p with an F is itself an F. This follows easily from the definition of the F's and the triangle inequality (9).

The direction numbers of the normal to S at a point of one of the curves F are $\partial z/\partial x$, $\partial z/\partial y$, -1, where the differentiation is performed in (12) and then y is eliminated by means of (13). Since $\nu \geqq p$ the direction numbers are

$$i + (p+1)c_{p+1}x^p + \cdots,\ d_\lambda x^\lambda + \cdots, -1,$$

except that in the middle number the leading term is not necessarily the one given, as a term in $x^{\nu+1}$ also may occur there. These direction numbers become i, 0, -1, the direction numbers of L, when $x = 0$, for L is the normal to S or T at O.

Let q stand for the order of coincidence of the normal of S with its limiting position L, evaluated along one of the curves F, relative to the parameter x. Applying (11) we find that

$$q = \min(\lambda, p). \tag{16}$$

Thus the purely geometric integer q depends on the surface S and the tangent line L, but not on the particular one of the curves F. It is just as intrinsic to the geometric object consisting of S and L as is p, the order of contact of S with L.

We now come to the general curve C on S, for which it is required to find the limit R of the ratio of arc to chord. If C is not tangent to L, R is 1. Assuming therefore that C is tangent to L, we may suppose it to have the equations (13) and (14), in which ν may again be any positive integer.

The order of contact n of C with L, defined by (15), has also another meaning. It represents what might be called the order of contact of C with the class of curves F; that is, the least order of contact which C has with any member of the class. The fact is that when $n < p$, C has contact of order n with all the F's, and when $n = p$ the orders of contact of C with the F's have all values from p upward, indeed C is itself one of the F's. These things are obvious either from the triangle inequality or from (13) and (14). Thus to assign the value of n, which is what we shall do in order to fix the position of C to a first approximation, has a geometric significance within the surface S. The intrinsic description of n is here less simple than in the non-isotropic tangent plane case, because there the curve of intersection D of S with the isotropic plane through L could be introduced as a reference curve, and here this cannot be done.

However, a portion of the intersection of S with the isotropic plane T does figure in the present case whenever $q \leqq \frac{1}{2}p$. The following lemma introduces it.

LEMMA. *In the class of regular analytic curves that lie on the surface S and have contact of order q or higher with the straight line L at O, there is none that lies in the tangent plane T when $q > \frac{1}{2}p$. When $q \leqq \frac{1}{2}p$, T contains exactly one such curve, G. The order of contact of G with L is $p - q$.*[6]

To prove this we observe from (12) that the whole intersection of S and T obeys the equation

$$(17) \qquad (c_{p+1}x^{p+1} + \cdots) + (d_\lambda x^\lambda + \cdots)y + (\cdots)y^2 + \cdots = 0.$$

To find a curve of the required kind lying in this intersection we substitute (13) with $\nu \geqq q$ into (17) in order to determine $\nu, a_{\nu+1}, a_{\nu+2}, \cdots$. If $q = p$, $\lambda \geqq p$ by (16), and obviously (17) cannot be satisfied. If $\frac{1}{2}p < q < p$, then $q = \lambda$ and $\nu \geqq \lambda$; as $\nu + \lambda + 1 > p + 1$, again no solution is possible. But if $q \leqq \frac{1}{2}p$, it is found that the power series (13) is uniquely determined with $\nu = p - \lambda$, that is $n = p - q$. The first $p - 2q$ coefficients $a_{\nu+1}, a_{\nu+2}, \cdots,$ $a_{2\nu-\lambda+1}$ (or $a_{p-q+1}, \cdots, a_{2p-3q+1}$) are obtained from

$$(18) \qquad \begin{aligned} &c_{p+1} + d_\lambda a_{\nu+1} = 0, \\ &c_{p+2} + d_\lambda a_{\nu+2} + d_{\lambda+1}a_{\nu+1} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\ &c_{2\nu+1} + d_\lambda a_{2\nu-\lambda+1} + \cdots + d_\nu a_{\nu+1} = 0. \end{aligned}$$

The question remains whether the power series thus determined is convergent. This may be settled as follows. In (17) substitute $y = ux^{p-\lambda+1}$ and remove the factor x^{p+1}. The result is

$$(19) \qquad \begin{aligned} &(c_{p+1} + c_{p+2}x + \cdots) \\ &\qquad + (d_\lambda + d_{\lambda+1}x + \cdots)u + (\cdots)u^2 + \cdots = 0, \end{aligned}$$

---

[6] The class of curves described exists only when $q \leqq p$, since no curve on S has contact of order greater than p with L. It can be shown that when $q < p$, $p - q$ is the order of coincidence with its limiting position L, along any curve F and relative to x, of that normal to F which touches S.

in which the coefficients of $u^2, u^3, \cdots$ all have the factor x. By the ordinary implicit function theorem there exists a convergent power series in x,

$$u = -\frac{c_{p+1}}{d_\lambda} + \cdots,$$

which satisfies (19). Multiplying it by $x^{p-\lambda+1}$ we obtain a convergent power series which satisfies (17). Since such a series exists, it must be the same as the unique one that was determined formally in the previous paragraph. This completes the proof of the lemma.

It is now possible to determine accurately enough for our problem the order of contact m between a curve C on S and the isotropic tangent plane T. At this point, however, three types of surfaces must be distinguished, depending on whether q is greater than, equal to, or less than $\frac{1}{2}p$. On a surface of the first type it is enough to know the value of n, the order of contact between C and L (or between C and the class of curves F), in order to decide about m. On surfaces of the second and third types it may be necessary also to be given the order of contact of C with the curve G defined in the lemma. This last order of contact we shall call k. By the triangle inequality $k \geqq \min(n, p - q)$, the equality sign certainly holding when $n \neq p - q$.

Surface of type I: $q > \frac{1}{2}p$. Under this type $\lambda > p/2$ by (16).

Case Iα: $n < \frac{1}{2}p$. From (15) $n = \nu$, and hence $\nu < p/2$. Thus $p + 1$, $\nu + \lambda + 1$ and $2\nu + 2$ all exceed $2n + 1$. It follows from (14) that $m > 2n$.

Cases Iβ: $n = \frac{1}{2}p$, *and Iγ:* $n > \frac{1}{2}p$. Then $\nu \geqq p/2$ and $2\nu + 2 > p + 1$. As $\lambda > p/2$, $\nu + \lambda + 1 > p + 1$. Therefore $m = p$.

Surface of type II: $q = \frac{1}{2}p$. Under this type $q = \lambda = p/2$.

Case IIα: $n < q$. As $n < p/2$, $n = \nu$ and $\nu < p/2$. Again $p + 1$, $\nu + \lambda + 1$ and $2\nu + 2$ all exceed $2n + 1$. Therefore $m > 2n$.

Case IIβ: $n = q$. As $n = \nu$, $p + 1 = \nu + \lambda + 1 < 2\nu + 2$. It was remarked above that $k \geqq n$ or q. Two subcases are necessary.

Subcase IIβ_1: $k = q$. The coefficient $a_{\nu+1}$ in (13) must have a different value for the curve C than for G. Hence the first of equations (18) does not hold. It is then clear from (14) that $m = p$.

Subcase IIβ_2: $k > q$. This time the first of equations (18) holds, and $m > p$.

Case IIγ: $n > q$. Then $\nu > q$. Hence $\nu + \lambda + 1 > p + 1$ and $2\nu + 2 > p + 1$. Therefore $m = p$.

Surface of type III: $q < \frac{1}{2}p$. Under this type $q = \lambda < p/2$.

Case IIIα: $n < q$. As in case *IIα*, $m > 2n$.

Cases IIIβ: $n = q$, and *IIIγ:* $q < n < p - q$. Then $n = \nu$, and $\nu + \lambda + 1 < p + 1$. Also $\nu + \lambda + 1 < 2\nu + 2$. Therefore $m = q + n$.

Case IIIδ: $n = p - q$. As $n = \nu$ and $\lambda < \nu$,

$$p + 1 = \nu + \lambda + 1 < 2\nu + 2.$$

Subcases IIIδ₁: $p - q \leqq k < 2p - 3q$, *and IIIδ₂:* $k = 2p - 3q$. The coefficients $a_{\nu+1}, \cdots, a_k$ in (13) must have the same values for the curve C as for G (this condition is vacuous when $k = p - q$), and it is easy to see that a_{k+1} must have different values. This means that of the equations (18) those beginning with $c_{p+1}, \cdots, c_{q+k}$ must hold for the curve C, while the one beginning with c_{q+k+1} must not. It is then clear from (14) that $m = q + k$.

Subcase IIIδ₃: $k > 2p - 3q$. Here equations (18) all hold. Therefore $m > 2n$.

Case IIIε: $n > p - q$. As in *IIγ*, $m = p$.

6. The isotropic curves on a surface. In Section 3 we made use of the two isotropic curves E and E' that pass through a point O of a surface S, when the tangent plane T at O is not an isotropic plane. The positions of these curves were described to the extent that their orders of contact with certain reference curves D and D' were given. This section is devoted to a similar discussion of the case when T is isotropic. The object therefore is *to describe the number and position of the regular analytic isotropic curves lying on a surface and passing through a point of the surface at which the tangent plane is isotropic.* The results are complete except for one case in which a question of convergence remains unsettled.

As in the previous section, L is the isotropic straight line in T, p is the order of contact of L with S, and q is the order of coincidence of the normal of S with L, evaluated along any curve on S of maximum contact with L.

The results are stated in the following three theorems, one for each of the three types of surfaces.

THEOREM III. *On a surface of type I* ($q > \frac{1}{2}p$), *when p is an even number there are two isotropic curves that pass through O. They have contact of order $\frac{1}{2}p$ with L and with one another. When p is odd there are no* (*regular analytic*) *isotropic curves.*

THEOREM IV. *On a surface of type II* ($q = \frac{1}{2}p$), *either there is just one isotropic curve, and it has contact of order q with L; or there are two, of contact q with L and with one another; or there is one isotropic curve of* *contact q with L, and there are (complex) ∞^1 other isotropic curves of contact q with L and with the first one, these ∞^1 having some single order of contact higher than q among themselves.*

THEOREM V. *On a surface of type III ($q < \frac{1}{2}p$), there are two isotropic curves through O. The order of contact of one of them with L is q, and of the other $p - q$.*

Whenever there are just two isotropic curves they will be called E_1 and E_2. On a surface of type *III*, E_1 will stand for the curve of contact q with L, and E_2 for the curve of contact $p - q$. On a surface of type *II*, if there is just one isotropic curve it will be called E_1. If there are $1 + \infty^1$, we shall refer to the single one as E_1 and to the ∞^1 as the E_2's.

To prove these theorems we take the equation of S in the form (12). If there exists a regular analytic isotropic curve on S passing through O, it has equations of the form (13), (14) obeying the condition

$$(20) \qquad 1 + y'^2 + z'^2 = 0.$$

On a surface of type *I*, $\lambda > p/2$ by (16). Suppose p is odd. From (13),

$$(21) \qquad y'^2 = (\nu + 1)^2 a^2_{\nu+1} x^{2\nu} + 2(\nu + 1)(\nu + 2) a_{\nu+1} a_{\nu+2} x^{2\nu+1} + \cdots.$$

If $\nu \geqq p/2$, then from (14),

$$(22) \qquad 1 + z'^2 = 2i(p + 1) c_{p+1} x^p + \cdots.$$

Since 2ν cannot equal p, the sum of (21) and (22) cannot obey (20). Similarly if $\nu < p/2$ it is easy to show that (20) cannot be obeyed. Hence when p is odd there are no isotropic curves on a surface of type *I*.

Suppose p is even. It may be shown first, that $\nu = p/2$; second, that $a_{\nu+1}$ must satisfy

$$(23) \qquad (\nu + 1)^2 a^2_{\nu+1} + 2i(p + 1) c_{p+1} = 0;$$

and third, that for each of the two distinct values of $a_{\nu+1}$ obeying (23) there exists a unique power series (13) such that (13) and (14) satisfy (20).

To ascertain that these series converge, we start from the differential equation of the isotropic curves on a surface whose equation is given in the form $z = f(x, y)$:

$$(24) \qquad (1 + f_y^2) y'^2 + 2 f_x f_y y' + (1 + f_x^2) = 0.$$

This may be written

$$(25) \qquad y' = \frac{-f_x f_y + [-(1 + f_x^2 + f_y^2)]^{\frac{1}{2}}}{1 + f_y^2}.$$

From (12),

$$f_x = [i + (p+1)c_{p+1}x^p + \cdots] + (\lambda d_\lambda x^{\lambda-1} + \cdots)y + (\cdots)y^2 + \cdots,$$
$$f_y = (d_\lambda x^\lambda + \cdots) + (\cdots)2y + \cdots.$$

Let

$$y = x^\nu u, \tag{26}$$

where $\nu = p/2$. As $\lambda > p/2$,

$$-(1 + f_x{}^2 + f_y{}^2) = x^{2\nu}[-2i(p+1)c_{p+1} + \cdots].$$

Since the power series in x and u within the brackets has a non-vanishing constant term, there are two distinct power series which are square roots of it. Substituting either one in (25) and dividing by $x^{\nu-1}$ we have

$$xu' + \nu u = \frac{x\{[-2i(p+1)c_{p+1}]^{\frac{1}{2}} + \cdots\}}{1 + \cdots}. \tag{27}$$

This differential equation is of the form

$$xu' - bu = a_{10}x + a_{20}x^2 + a_{11}xu + a_{02}u^2 + \cdots, \tag{28}$$

where $b = -\nu$ and $a_{10} = [-2i(p+1)c_{p+1}]^{\frac{1}{2}}$.

Briot and Bouquet proved that when b is not a positive integer, (28) has a power series solution

$$u = \frac{a_{10}}{1-b}x + \cdots$$

which vanishes when $x = 0$, and only one such solution.[7] Consequently there exist two convergent power series,

$$y = \frac{\pm[-2i(p+1)c_{p+1}]^{\frac{1}{2}}}{1+\nu}x^{\nu+1} + \cdots,$$

that obey (24) or (20). These must be the series that were determined formally above. Theorem III is therefore proved.

In the case of Theorem V, $q = \lambda < p/2$. It can be shown as above, though the details are more complicated, that for an isotropic curve either $\nu = \lambda$ or $\nu = p - \lambda$, and that corresponding to either value of ν just one power series (13) can be formally determined. When $\nu = p - \lambda$ the curve E_2 has contact of order $k = 2p - 3q$ with the curve G defined in Section **5**.

In the convergence proofs we make instead of (26) the substitution

$$y = x^{\nu+1}(a_{\nu+1} + u), \tag{29}$$

where $a_{\nu+1}$ is the leading coefficient of the power series formally determined in the preceding paragraph. (It is convenient to do this in two steps, first $y = x^{\nu+1}v$, then $v = a_{\nu+1} + u$.) After this substitution has been completed and x^{ν} has been divided out, two distinct square roots may be extracted as before. By choosing a suitable one of the two roots, we can reduce the equation to the form (28) with b equal to $-(\lambda+1)(2\lambda+1)/(3\lambda+1)$ in the case $\nu=\lambda$; and with more difficulty to the same form with b equal to $-p-1$ in the case $\nu=p-\lambda$.

---

[7] *Journal de l'École Polytechnique*, vol. XXI (1856); or Goursat, *Cours d'Analyse Mathématique*, 4th ed. (1924), vol. II, pp. 523-526. The writer owes this reference to Professor Kasner.

Finally consider a surface (12) of type *II* $(q=\lambda=p/2)$. In order for the curve (13), (14) to be isotropic we find that ν must equal λ and $a_{\nu+1}$ must be a root of the quadratic equation

$$(30)\qquad (\lambda+1)^2a^2_{\nu+1}+2i(p+1)d_\lambda a_{\nu+1}+2i(p+1)c_{p+1}=0.$$

There are four mutually exclusive special subclasses of surfaces of type *II* of which account has to be taken. Two of these are defined respectively by the conditions

$$(31)\qquad 2i(\lambda+1)^2c_{p+1}+(p+1)d_\lambda^2=0,$$

$$(32)\qquad 2i(p+1)(\lambda+1)c_{p+1}+(p+\lambda+1)d_\lambda^2=0.$$

The third subclass consists of those surfaces of type *II* for which there exists some integer $\theta>1$ such that

$$(33)\qquad 2i(p+1)(\lambda+1)(\lambda+\theta)^2c_{p+1} + (p+\theta)[p\lambda+(p+\lambda+1)\theta]d_\lambda^2=0.$$

The fourth subclass is defined by $p=q=\infty$, and (30) does not apply to it.

A general surface of type *II*, that is one which belongs to none of the four subclasses, contains two distinct isotropic curves of contact $p/2$ with L and with one another. For since (31) does not hold, (30) has two distinct roots $a_{\nu+1}$. Since (33) does not hold, it can be shown that corresponding to each root a unique series (13) can be determined such that the curve (13), (14) is isotropic. In the convergence proof, using the substitution (29) and then employing (30) we arrive at an equation

$$(34)\qquad xu'+(\lambda+1)(u+a_{\nu+1})= \frac{-id_\lambda+x(\cdot\cdot\cdot)+[[(\lambda+1)a_{\nu+1}+id_\lambda]^2-2i\lambda d_\lambda u+x(\cdot\cdot\cdot)]^{\frac{1}{2}}}{1+x^p(\cdot\cdot\cdot)}.$$

Since (32) does not hold, $(\lambda+1)a_{\nu+1}+id_\lambda$ cannot vanish for either root $a_{\nu+1}$. By a suitable choice of the square root, (34) can be reduced to the form (28) with

$$b=-\frac{(\lambda+1)^2a_{\nu+1}+(p+1)id_\lambda}{(\lambda+1)a_{\nu+1}+id_\lambda}.$$

This b cannot equal a positive integer $\theta-1$, where θ is an integer greater than 1, or (33) would hold. Hence there are just two isotropic curves, E_1 and E_2.

For the subclass of surfaces defined by (31) the quadratic equation (30) has a double root. In this case it can be shown just as above that there is a unique isotropic curve E_1.

For the subclass (32) it is possible to show as in the general case that just two distinct power series (13) can be formally determined. Convergence can be proved as above for one of these, so that there is certainly at least one isotropic curve E_1. The $a_{\nu+1}$ for the other series (13) obeys the relation

$$(\lambda+1)a_{\nu+1}+id_\lambda=0. \tag{35}$$

On account of this the constant term of the series under the square root sign in (34) vanishes, and we cannot proceed in the same way as before. The writer does not know whether this second series converges or not, and hence does not know whether or not there is a second isotropic curve E_2 in this case.

However, for the purpose of our main problem there is a way around the difficulty. It can be shown that on a surface of type *II* the locus of points at which the tangent plane to the surface is isotropic includes one curve and only one which has contact of order q with L. The proof is similar to that of the lemma in Section **5**. Let this curve be called W. Now the $a_{\nu+1}$ of W obeys (35). Consequently if there is an E_2, W has contact with it of order higher than q. In the problem of the limit of the ratio of arc to chord, all that matters about E_2 is its leading coefficient $a_{\nu+1}$. Hence for that problem W, having the same $a_{\nu+1}$ as E_2, will serve instead.

For the subclass (33) there are two distinct values of $a_{\nu+1}$. Corresponding to one of these it may be proved as above there exists an isotropic curve E_1. Corresponding to the other $a_{\nu+1}$ the formal determination of the remaining coefficients in (13) comes to a stop at $a_{\nu+\theta}$. In general no $a_{\nu+\theta}$ and hence no curve E_2 exists. Only if a certain relation among $c_{p+\theta}$ and other coefficients in (12) is satisfied can $a_{\nu+\theta}$ be determined, and then in fact it can be taken arbitrarily. The determination of the further coefficients in (13) with any given choice of $a_{\nu+\theta}$ is then possible and unique. In the convergence proof an equation of the form (28) is obtained with b equal to $\theta-1$, a positive integer. In this case it follows from the work of Briot and Bouquet that when the series (13) can be determined formally, it converges. Consequently either there is no E_2; or else there are ∞^1 E_2's corresponding to an arbitrary complex parameter $a_{\nu+\theta}$, having contact of order q with E_1 and $q+\theta-1$ with one another. When there is no E_2 let a curve U be drawn on the surface with ν equal to q, $a_{\nu+1}$ equal to the value for which E_2 could not be determined, and $a_{\nu+2}, a_{\nu+3}, \cdots$ chosen in any manner so that (13) converges.

When $p=q=\infty$ the surface contains a single isotropic curve E_1, namely the isotropic straight line L. This last subclass of type *II*, it can be shown, includes all isotropic developable surfaces.

The following statements can be proved by simple computations that we omit. Among the curves on a surface of type I which have contact of order $p/2$ with L, those and only those have contact higher than p with some isotropic curve (not necessarily on the surface), which have contact higher than $p/2$ with E_1 or else with E_2. On a surface of type II curves of contact q with L have contact higher than p with an isotropic curve when, and only when, they have contact higher than q with E_1 or E_2, in the general case of just two isotropic curves; with E_1, in the case of the subclass (31); with E_1 or W in the case of the subclass (32); with E_1 or any of the E_2's or U in the case of the subclass (33). On a surface of type III curves of contact q with L have contact higher than $2q$ with an isotropic curve when, and only when, their contact with E_1 is higher than q; and curves of contact $p-q$ with L and $2p-3q$ with G have contact higher than $2(p-q)$ with an isotropic curve when, and only when, their contact with E_2 is higher than $2p-3q$.

7. The limit at an isotropic point of a surface. A curve C on a surface S passes through a point O of S at which the tangent plane T is isotropic. What is the limit R of the ratio of arc to chord of C at O? The answer follows from Section **6**, the table at the end of Section **5**, and Theorem II, and we need only state what it is.

To summarize, we recall that L is the single isotropic straight line in T that passes through O, p is its order of contact with S, F is any curve on S having contact of order p with L, and q is the order of coincidence of the normal of S with L evaluated along any F. The order of contact of C with L or with the F's is denoted by n. The tables below are arranged according to ascending values of n, from $n=0$, when C passes through O in a direction other than that of L, to the maximum $n=p$. The abbreviation R_h was defined in Section **3**. Even though C is confined to S, it can be shown that δ, the derivative of the radius of curvature with respect to arc-length, is arbitrary enough so that the formulas in which it appears below can assume all the values indicated. On surfaces of types I and III, E_1 and E_2 are the isotropic curves of Theorems III and V. On surfaces of types II and III, G is a curve of intersection of S and T described in Section **5**, and k stands for the order of contact CG. The isotropic curves E_1 and E_2, the isotropic tangent plane locus W, and the curve U on a surface of type II were defined in Section **6**.

Surface of type I: $q > \frac{1}{2}p$

$$n = 0 : R = 1$$

$$0 < n < p/2 : R = \pm 1$$

$$n = p/2,\ CE_1 = CE_2 = p/2 : R = \delta/[\delta^2 - 1/R_p{}^2]^{\frac{1}{2}},$$

which may be any number but ± 1, R_p and 0,

$$\begin{aligned} n = p/2,\ CE_1 \text{ or } CE_2 > p/2 &: R = 0 \\ p/2 < n \leqq p &: R = R_p. \end{aligned}$$

Surface of type II: $q = \frac{1}{2}p$

$$\begin{aligned} n = 0 &: R = 1 \\ 0 < n < q &: R = \pm 1 \\ n = q,\ k = q,\ CE_1 = CE_2 = CW = CU = q &: R = \delta/[\delta^2 - 1/R_p{}^2]^{\frac{1}{2}}, \end{aligned}$$

which may be any number but ± 1, R_p and 0 (if there is an E_2, W or U),

$$\begin{aligned} n = q,\ k = q,\ CE_1 \text{ or } CE_2 \text{ or } CW \text{ or } CU > q &: R = 0 \\ n = q,\ k > q &: R = \pm 1 \\ q < n \leqq p &: R = R_p. \end{aligned}$$

Surface of type III: $q < \frac{1}{2}p$

$$\begin{aligned} n = 0 &: R = 1 \\ 0 < n < q &: R = \pm 1 \\ n = q,\ CE_1 = q &: R = \delta/[\delta^2 - 1/R^2{}_{2q}]^{\frac{1}{2}}, \end{aligned}$$

which may be any number but ± 1, R_{2q} and 0,

$$\begin{aligned} n = q,\ CE_1 > q &: R = 0 \\ q < n < p - q &: R = R_{q+n} \\ n = p - q,\ p - q \leqq k < 2p - 3q &: R = R_{q+k} \\ n = p - q,\ k = 2p - 3q,\ CE_2 = 2p - 3q &: R = \delta/[\delta^2 - 1/R^2{}_{2(p-q)}]^{\frac{1}{2}}, \end{aligned}$$

which may be any number but ± 1, $R_{2(p-q)}$ and 0,

$$\begin{aligned} n = p - q,\ k = 2p - 3q,\ CE_2 > 2p - 3q &: R = 0 \\ n = p - q,\ k > 2p - 3q &: R = \pm 1 \\ p - q < n \leqq p &: R = R_p. \end{aligned}$$

Under type *I*, if p is odd the third and fourth cases do not occur; if $p = 2$ the second case does not occur. Under type *II*, if $p = 2$ the second case does not occur; if $p = q = \infty$ only the first two cases occur. Under type *III*, if $q = 1$ the second case does not occur; if $p = 2q + 1$ the fifth case does not occur; if $p = \infty$ only the first five cases occur. Theorem VIII of Kasner's 1932 paper, dealing with the various possible totalities of limits at an isotropic point of a surface, follows from this enumeration.

The table shows that the limit always begins with the value 1, and ends with R_p when the curve is, so to speak, in the shadow of L. But from first to last the distinct stages, if the surface is of type *III*, can number as many as ten.

St. John's College,
Annapolis, Md.

SEMI-ANALYTIC UNITARY SUBSPACES OF UNITARY SPACE.*

By N. Coburn.

I. Introduction. We consider an n-dimensional unitary space K_n (real topological space of $2n$ dimensions) as the imbedding manifold. Into this K_n, we imbed an Hermitian X_m of m dimensions ($m < n$; real topological space of $2m$ dimensions). Since the theory of analytic unitary subspaces K_m in K_n has been investigated,[1] we restrict ourselves to the theory of semi-analytic unitary subspaces in K_n. We denote such a subspace by S_m.

Our approach is from the group viewpoint. By use of composite invariants of K_n, we construct a theory of X_m in K_n which is a generalization of the theory of K_m in K_n. It is shown that use of non-composite invariants of K_n leads to a more general but very complicated theory of X_m in K_n. Now, the X_m is invariant under the semi-analytic group of subspace coördinate transformations. By definition, X_m is an S_m when a set of preferred coördinate systems, related by a group of coördinate transformations (the group of S_m), exists for which certain components of the connection of X_m vanish. It is shown that the analytic group is a subgroup of the group of S_m which in turn is a subgroup of the semi-analytic group of X_m. The metric tensor of S_m is discussed and it is shown that if a coördinate system exists in S_m for which the ranks of certain matrices formed from the metric tensor of S_m are m, *then* S_m *is a unitary Euclidean space.* This is a surprising result. Analytically, it means that the conditions that an Hermitian X_m be an S_m are very restrictive.

II. Composite affinors of $\mathbf{X}_m$ and $\mathbf{K}_n$. Let the coördinates of K_n be denoted by the complex numbers ξ^λ ($\lambda, \mu = 1, 2 \cdot \cdot \cdot n$). In terms of their real and imaginary parts, these numbers are given by

$$(2.1) \qquad \xi^\lambda = x^\lambda + iy^\lambda, \qquad i = \sqrt{-1}.$$

For the conjugate variable $\xi^{\lambda *}$, we write

$$(2.2) \qquad \xi^{\lambda *} = x^\lambda - iy^\lambda.$$

Thus, we may associate with any vector function of position $v^\lambda(\xi^\lambda, \xi^{\lambda *})$, the conjugate vector $v^{\lambda *}(\xi^{\lambda *}, \xi^\lambda)$. This latter vector is obtained by replacing i by $-i$ wherever i occurs in the former vector. In particular ξ^λ is replaced by $\xi^{\lambda *}$ and vice versa. In the case of an affinor, the conjugate affinor is formed

---

* Received June 2, 1941.

in the same manner as indicated above for a vector. This conjugate affinor is indicated by starring each unstarred index of the original affinor and removing the star from each starred index of the original affinor. Thus, if $a_{\lambda\mu^*}$ denotes the metric tensor of K_n, we write[2]

$$(2.3) \qquad (a_{\lambda\mu^*})^* = a_{\lambda^*\mu}.$$

It can be easily shown that the existence of any affinor equation implies the existence of a conjugate affinor equation which is obtained by replacing i by $-i$ and any unstarred index by the corresponding starred index or vice versa. We denote this result by writing "conj." to the right of the original equation.

Another operation which we use, is the formation of the transpose. This term is borrowed from matrix notation and is only applicable to affinors of the second order. It implies that two indices are to be interchanged; the star of an index retains its position. Thus,

$$(2.4) \qquad (a_{\lambda\mu^*})' = a_{\mu\lambda^*}.$$

We now require that the metric tensor of K_n shall possess Hermitian symmetry. By definition, this means that

$$(2.5) \qquad [(a_{\lambda\mu^*})^*]' = a_{\lambda\mu^*}.$$

By use of the previous definitions, (2.5) becomes

$$(2.6) \qquad a_{\lambda\mu^*} = a_{\mu^*\lambda}.$$

A simple example will show that (2.5) or (2.6) implies that the off-diagonal elements of the matrix $a_{\lambda\mu^*}$ are Hermitian symmetric. Thus, if $a_{12^*} = 3 - 7i$, then $a_{21^*} = 3 + 7i$, etc.:

We conclude the discussion of K_n by defining its transformation group.

Definition 1. The allowable coördinate transformations in K_n shall be analytic with non-vanishing Jacobian

$$(2.7) \qquad \xi'^{\lambda} = \xi'^{\lambda}(\xi^{\lambda}), \quad \text{conj.}$$

In most of our work, we shall find it advantageous not to distinguish between starred and unstarred indices of K_n. Thus, we shall use the single Greek capital, say Λ, to represent the corresponding Greek small case letter and its star, say λ, λ^*. For instance, we write

$$(2.8) \qquad \Lambda = \lambda, \lambda^*; \quad \mathrm{M} = \mu, \mu^*; \quad \mathrm{P} = \rho, \rho^*; \quad \Sigma = \sigma, \sigma^*; \quad \mathrm{T} = \tau, \tau^*; \text{ etc.}$$

Note that there are now three invariants under the group (2.7). For, if $v^{\Lambda} \to (v^{\lambda}, v^{\lambda^*})$ and $w_{\Lambda} \to (w_{\lambda}, w_{\lambda^*})$ are two vectors of K_n, then

$$w_\lambda v^\lambda; \quad w_{\lambda^*} v^{\lambda^*}; \quad w_\Lambda v^\Lambda, \tag{2.9}$$

are the three invariants under (2.7). Of these three, the last is a composite invariant, being the sum of the first two. Hence, use of the Greek capital letters is restrictive in that only composite invariants enter. On the other hand, use of the lower case Greek letters requires the introduction of several affinors of the same type. This point will be stressed in (2.31) to (2.38) when certain of the projection-affinor's components are discussed. Choosing between these evils, we elect to use the Greek capitals and hence composite invariants of K_n.

Let the equations of the semi-analytic X_m in K_n be given by

$$\xi^\Lambda = \xi^\Lambda(\eta^a, \eta^{a^*}), \quad \text{conj.}, \qquad a, b, c, d = 1, 2, \cdots m, \tag{2.10}$$

where η^a, η^{a^*} are the complex surface coördinates of X_m and $m < n$. These surface coördinates can be expressed in terms of real variables u^a, v^a by means of

$$\eta^a = u^a + iv^a; \qquad \eta^{a^*} = u^a - iv^a. \tag{2.11}$$

In a manner similar to (2.8), we write

$$A = a, a^*; \quad B = b, b^*; \quad C = c, c^*; \quad D = d, d^*; \text{ etc.} \tag{2.12}$$

However, in this case, the notation is forced upon us. This can be seen by defining the group of allowable coördinate transformations in X_m;

Definition 2. The allowable coördinate transformations in X_m shall be semi-analytic with non vanishing Jacobian

$$\eta'^a = \eta'^a(\eta^a, \eta^{a^*}), \quad \text{conj.} \tag{2.13}$$

Evidently, such transformations form a group. We observe that if $w_A \rightarrow (w_a, w_{a^*})$ and $v^A \rightarrow (v^a, v^{a^*})$ are vectors of X_m, then the single invariant under (2.13) is $w_A v^A$. Note that neither $w_a v^a$, nor $w_{a^*} v^{a^*}$, is individually invariant; their sum $w_A v^A$ is invariant. Hence, for the present, we shall confine ourselves to quantities of X_m whose components are expressible in large type indices.

The projection affinor of X_m in K_n is obtained by forming the various partial derivatives of (2.10). We write

$$B_a{}^\lambda = \partial_a \xi^\lambda, \qquad B_{a^*}{}^\lambda = \partial_{a^*} \xi^\lambda, \quad \text{conj.}, \qquad \partial_a = \partial/\partial\eta^a. \tag{2.14}$$

By use of (2.8), (2.12), we may write the above in the concise form

$$B_A{}^\Lambda = \partial_A \xi^\Lambda. \tag{2.15}$$

The affinor (2.15) can be considered as being composed of $2m$ vectors of K_n

which span the local tangent Euclidean unitary space U_m of X_m. By forming the differential of (2. 10), we find

$$d\xi^\Lambda = B_A{}^\Lambda d\eta^A. \tag{2.16}$$

For the element of arc length in K_n, we have the well known formula

$$ds^2 = a_{\lambda\mu^*} d\xi^\lambda d\xi^{\mu^*}. \tag{2.17}$$

With the aid of (2. 6), we may write the above formula in the more symmetric form

$$2ds^2 = (a_{\lambda\mu^*} + a_{\mu^*\lambda}) d\xi^\lambda d\xi^{\mu^*}, \tag{2.18}$$

and by use of (2. 8), the equation (2. 18) becomes

$$2ds^2 = a_{\Lambda M} d\xi^\Lambda d\xi^M, \qquad a_{\lambda\mu} = a_{\lambda^*\mu^*} = 0. \tag{2.19}$$

Substituting (2. 16) into (2. 19), we obtain the element of arc length along X_m,

$$2ds^2 = a_{\Lambda M} B_C{}^\Lambda B_D{}^M \, d\eta^C \, d\eta^D. \tag{2.20}$$

We define the metric tensor of X_m as the X_m projection of the metric tensor $a_{\Lambda M}$ (of rank $2n$) of K_n,

$$a_{CD} = a_{\Lambda M} B_C{}^\Lambda B_D{}^M. \tag{2.21}$$

Evidently, the rank r of a_{CD} is $r \leq 2m$. We shall require that X_m possess a metric tensor a_{CD} of rank $r = 2m$.

We briefly discuss some properties of a_{CD}. By substituting (2. 21) into (2. 20), we obtain the formula for X_m arc length

$$2ds^2 = a_{CD} \, d\eta^C \, d\eta^D. \tag{2.22}$$

From (2. 6), we see that $a_{\Lambda M}$ is symmetric (formally). Hence from (2. 21), it follows that

$$a_{CD} = a_{DC}. \tag{2.23}$$

Furthermore, by writing out the relations (2. 21) for specific C, D, and forming the conjugate equations, we obtain the following relations which are invariant under (2. 13),

$$(a_{cd})^* = a_{c^*d^*}, \quad (a_{cd^*})^* = a_{c^*d}, \quad \text{conj.} \tag{2.24}$$

Note that the relations (2. 23) are equivalent to

$$a_{cd} = a_{dc}, \qquad a_{cd^*} = a_{d^*c}, \qquad a_{c^*d^*} = a_{d^*c^*}. \tag{2.25}$$

By forming the transpose of the second equation in (2. 24) and its conjugate and use of (2. 25), we obtain

$$[(a_{cd^*})^*]' = a_{cd^*}, \qquad [(a_{c^*d})^*]' = a_{c^*d}. \tag{2.26}$$

We summarize these results as follows:

THEOREM 1. *The components of the metric tensor of* X_m, $a_{CD} \rightarrow (a_{cd}, a_{cd^*}, a_{c^*d}, a_{c^*d^*})$ *have the following properties* (1) *the matrices* a_{cd^*}, a_{c^*d} *are each Hermitian symmetric;* (2) *if we order the rows of the matrices* a_{cd^*}, a_{c^*d} *by the first index, then the* a_{cd^*} *matrix is the transpose of the* a_{c^*d} *matrix and vice versa;* (3) *the matrices* a_{cd}, $a_{c^*d^*}$ *are each real symmetric;* (4) *if we order the rows of these matrices as in* (2), *then* a_{cd} *is conjugate to* $a_{c^*d^*}$ *and vice versa.*

Let us define the contravariant components (a^{DC}) of the metric tensor of X_m by means of the equivalent formulas

$$a^{DC}a_{ED} = B_E^{C}, \qquad a^{CD}a_{DE} = B_E^{C}, \tag{2.27}$$

where B_E^{C} represents the unit affinor of X_m whose components have the values 0 (for $C \neq E$) and 1 (for $C = E$). These components (a^{DC}) satisfy relations of the form (2.24), (2.25).

Finally, we define the remaining components of the projection affinor of X_m in K_n. By analogy with the theory of the analytic K_m in K_n, we write

$$B_{\Lambda}^{\cdot C} = a^{DC}B_D^{M}a_{\Lambda M}. \tag{2.28}$$

By transvecting (2.28) with B_E^{Λ} and using (2.21), (2.27), we obtain

$$B_E^{\Lambda}B_{\Lambda}^{C} = B_E^{C}. \tag{2.29}$$

We conclude this section by defining the components

$$B_M^{\Lambda} = B_D^{\Lambda}B_M^{D}. \tag{2.30}$$

We are now in a position to show the difficulties attending the use of non-composite affinors of K_n. The equations (2.15), (2.16) must now be written as

$$B_A^{\lambda} = \partial_A\xi^{\lambda}, \quad \text{conj.} \tag{2.31}$$

$$d\xi^{\lambda} = B_A^{\lambda}d\eta^{A}, \quad \text{conj.} \tag{2.32}$$

By virtue of (2.32), the relation (2.17) becomes

$$ds^2 = a_{\lambda\mu^*}B_C^{\lambda}B_D^{\mu^*}\, d\eta^{C}\, d\eta^{D}. \tag{2.33}$$

The vital question arises now: "How shall we define the metric tensor of X_m?" An apparent clue is the fact that only the symmetric part of $a_{\lambda\mu^*}B_C^{\lambda}B_D^{\mu^*}$ enters into (2.33). Let us try the definition

$$a_{CD} = a_{\lambda\mu^*}B_{(C}^{\lambda}B_{D)}^{\mu^*}, \tag{2.34}$$

where the parentheses indicate a symmetric product formed on the indices involved. The corresponding definition of (2.28) would be

$$B_{\lambda}{}^{C} = a^{DC} B_{D}{}^{\mu^*} a_{\lambda\mu^*}, \quad \text{conj.} \tag{2.35}$$

But then, the result corresponding to (2.29) is no longer valid and this result is essential for some of the remaining theory. The only other apparent possibility is to define

$$a_{CD} = a_{\lambda\mu^*} B_{C}{}^{\lambda} B_{D}{}^{\mu^*}. \tag{2.36}$$

The metric affinor a_{CD} is now no longer symmetric. The results corresponding to (2.29) become

$$B_{\lambda}{}^{C} B^{\lambda}{}_{E} = a^{DC} B_{D}{}^{\mu^*} B_{E}{}^{\lambda} a_{\lambda\mu^*} = a^{DC} a_{ED} = B_{E}{}^{C}. \tag{2.37}$$

$$B_{\lambda^*}{}^{C} B_{E}{}^{\lambda^*} = a^{DC} B_{D}{}^{\mu} B_{E}{}^{\lambda^*} a_{\mu\lambda^*} = a^{DC} a_{DE} \neq B_{E}{}^{C}. \tag{2.38}$$

Since (2.37), (2.38) should be conjugates, the theory must be modified. These difficulties are responsible for our decision to use composite affinors of K_n.

III. The connection of X_m in K_n. First, we shall study the X_m covariant differential of X_m affinors. We define such a differential as the X_m projection of the K_n covariant differential, that is,

$$\delta' v^C = B_{\Lambda}{}^{C} \delta v^{\Lambda}, \qquad \delta' w_C = B_{C}{}^{\Lambda} \delta w_{\Lambda}. \tag{3.1}$$

Next, we define the X_m covariant derivative of an X_m affinor as the X_m projection of the K_n covariant derivative of the affinor, that is,

$$\begin{aligned} \nabla'_D v^C &= B_D{}^{\Lambda} B_M{}^{C} \nabla_{\Lambda} v^M, \\ \nabla'_D w_C &= B_D{}^{\Lambda} B_C{}^{M} \nabla_{\Lambda} w_M. \end{aligned} \tag{3.2}$$

By multiplying (3.2) with $d\eta^D$ and use of (2.16), (3.1), we obtain

$$\begin{aligned} \delta' v^C &= d\eta^D \nabla'_D v^C, \\ \delta' w_C &= d\eta^D \nabla'_D w_C. \end{aligned} \tag{3.3}$$

The development of the theory is evidently analogous to that of Riemannian space save for the fact that the indices have a duel range.[3] In particular, we note that (3.1) may be written as

$$\delta' v^C = B_{\Lambda}{}^{C} (dv^{\Lambda} + \Gamma^{\Lambda}{}_{MP}\, d\xi^M\, v^P), \tag{3.4}$$

where $\Gamma^{\Lambda}{}_{MP}$, the connection of K_n, has the non-vanishing components $(\Gamma^{\lambda}{}_{\mu\rho}, \Gamma^{\lambda^*}{}_{\mu^*\rho^*})$ and

$$\Gamma^{\Lambda}{}_{MP} = (\partial_M a_{PT}) a^{T\Lambda}. \tag{3.5}$$

Now (3.4) may be written in either of the two following forms;

$$\delta' v^C = dv^C - v^D d\eta^E B^\Lambda{}_D \partial_E B_\Lambda{}^C + B^{CM\Lambda}_{PED} \Gamma^P{}_{M\Lambda} v^D d\eta^E, \tag{3.6}$$

$$\delta' v^C = dv^C + \Gamma^C{}_{ED} v^D d\eta^E, \tag{3.7}$$

where $\Gamma^C{}_{ED}$ are the components of the X_m connection. Hence, by comparing the right hand sides of (3.6), (3.7), we obtain

$$\Gamma^C{}_{ED} = B^{CM\Lambda}_{PED} \Gamma^P{}_{M\Lambda} + B^C{}_\Lambda \partial_E B^\Lambda{}_D, \qquad B^{CM\Lambda}_{PED} = B^C{}_P B^M{}_E B^\Lambda{}_D. \tag{3.8}$$

We now require that the X_m shall be a semi-analytic unitary space S_m. This means that the connection of X_m shall have as its non vanishing components $(\Gamma^c{}_{ed}, \Gamma^{c^*}{}_{e^*d^*})$. But this last statement has no meaning unless a set of preferred coördinate systems is specified in X_m (the $\Gamma^C{}_{ED}$ are not the components of a tensor under the group 2.13). Furthermore, the group of S_m must transform one preferred coördinate system into another such coördinate system. With this in mind, we write

Definition 3. An X_n shall be denoted by S_m if: (1) a set of preferred coördinate systems exist in X_m such that

$$\Gamma^c{}_{e^*d} = \Gamma^c{}_{ed^*} = \Gamma^{c^*}{}_{ed} = 0, \quad \text{conj.}, \tag{3.9}$$

(2) the coödinate transformations from one set of preferred coördinates to another such set form a group.

We shall refer to the group entering the above definition as the group of S_m. By the above definition, the characteristic property of the group of S_m is that the $\Gamma^c{}_{e^*d}$, $\Gamma^c{}_{ed^*}$, $\Gamma^{c^*}{}_{ed}$, transform under this group as tensors. It is difficult to determine this group of S_m. However, we shall prove

THEOREM 2. *The group of analytic transformations*

$$\eta'^a = \eta'^a(\eta^a), \quad \text{conj.}, \tag{3.10}$$

is a subgroup of the group of S_m.

By expanding (3.8), we obtain

$$\begin{aligned} \Gamma^c{}_{e^*d} &= B^{cM\Lambda}_{Pe^*d} \Gamma^P_{M\Lambda} + B^c{}_\Lambda \partial_{e^*} B^\Lambda{}_d, \quad \text{conj.}, \\ \Gamma^c{}_{de^*} &= B^{cM\Lambda}_{Pde^*} \Gamma^P_{M\Lambda} + B^c{}_\Lambda \partial_d B^\Lambda{}_{e^*}, \quad \text{conj.}, \\ \Gamma^{c^*}{}_{de} &= B^{c^*M\Lambda}_{Pde} \Gamma^P_{M\Lambda} + B^{c^*}{}_\Lambda \partial_e B^\Lambda{}_d, \quad \text{conj.}, \\ \Gamma^c{}_{ed} &= B^{cM\Lambda}_{Ped} \Gamma^P_{M\Lambda} + B^c{}_\Lambda \partial_e B^\Lambda{}_d, \quad \text{conj.} \end{aligned} \tag{3.11}$$

By differentiation of (3.10), we obtain the equations

$$B'_c{}^c = \partial_c \eta'^c, \qquad B^d{}_{'d} = \partial_{'d} \eta^d, \quad \text{conj.} \tag{3.12}$$

$$(3.13) \qquad \partial_{'e^*}B^{d}{}_{'d} = 0, \qquad \partial_{e^*}B'^{e}{}_{c} = 0, \quad \text{conj.}$$

From (3.10), (2.15) and (2.29), we can easily prove

$$(3.14) \qquad B^{\Lambda}{}_{'c} = B^{\Lambda}{}_{c}B^{c}{}_{'c}, \qquad B'^{c}{}_{\Lambda} = B^{c}{}_{\Lambda}B'^{c}{}_{c}, \quad \text{conj.}$$

Multiplying the first equation of (3.11) by $B'^{\,c e^* d}_{\,c' e^{*\prime} d}$ and using (3.14), (3.13), the equation (3.11) becomes

$$(3.15) \qquad B'^{\,c e^* d}_{\,c' e^{*\prime} d}\ \Gamma^{c}{}_{e^* d} = \Gamma'^{c}{}_{'e^*'d}, \quad \text{conj.}$$

where

$$(3.16) \qquad \Gamma'^{c}{}_{'e^*'d} = B'^{\,c M\Lambda}_{\,P' e^{*\prime} d}\,\Gamma^{P}_{M\Lambda} + B'^{c}{}_{\Lambda}\,\partial_{'e^*}B^{\Lambda}{}_{'d}, \quad \text{conj.}$$

The second equation of (3.11) furnishes a similar result. However, the third equation is more troublesome. Upon multiplying by $B'^{\,c^* d e}_{\,c^{*\prime} d' e}$ we obtain

$$(3.17) \qquad B'^{\,c^* d e}_{\,c^{*\prime} d' e}\ \Gamma^{c^*}{}_{de} = B'^{\,c^* M\Lambda}_{\,P' d' e}\,\Gamma^{P}_{M\Lambda} + B'^{\,c^* e}_{\,\Lambda' e}\,\partial_{'d}B^{\Lambda}{}_{e}, \quad \text{conj.}$$

The second term on the right hand side of (3.17) may be written as

$$(3.18) \qquad B'^{c^*}{}_{\Lambda}\,\partial_{'d}B^{\Lambda}{}_{'e} - B'^{c^*}{}_{\Lambda}B^{\Lambda}{}_{e}\,\partial_{'d}B^{e}{}_{'e}.$$

With the aid of the formula (2.29), we find

$$(3.18) \qquad B'^{c^*}{}_{\Lambda}B^{\Lambda}{}_{e} \doteq B'^{c^*}{}_{c^*}(B^{c^*}{}_{\Lambda}B^{\Lambda}{}_{e}) = 0.$$

Hence the second term of (3.18) vanishes. Substituting (3.18) into (3.17), the right hand side of the latter is, by definition, $\Gamma'^{c^*}{}_{'d'e}$ and hence (3.17) becomes

$$(3.20) \qquad B'^{\,c^* d e}_{\,c^{*\prime} d' e}\,\Gamma^{c^*}{}_{de} = \Gamma'^{c^*}{}_{'d'e}, \quad \text{conj.}$$

The equations (3.15), (3.20) establish our theorem since (3.9) has been shown to be an invariant equation under (3.10). Since (3.9) is not invariant under the group of semi-analytic transformations (2.13), the group of S_m is a subgroup of this group (2.13). Hence, we have found upper and lower "bounds" for the group of S_m.

IV. Properties of the metric tensor of S_m. If we denote the metric tensor of S_m in K_n coördinates by $a'_{\Lambda M}$, then from (2.21), (3.1), we obtain

$$(4.1) \qquad a'_{\Lambda M} = B^{P\Sigma}_{\Lambda M}\,a_{P\Sigma}; \qquad \delta' a_{CE} = B^{\Lambda M}_{CE}\,\delta a'_{\Lambda M}.$$

Since $a_{\Lambda M}$ is the fundamental tensor of K_n, it has the property

$$(4.2) \qquad \delta a_{\Lambda M} = 0.$$

By taking the covariant differential of the first equation in (4.1) and using (4.2), the second equation in (4.1) becomes

$$\delta' a_{CB} = 0. \tag{4.3}$$

Hence the metric tensor a_{CB} of S_m is also a fundamental tensor. We conclude this introductory section by defining the Riemann-Christoffel affinor of S_m,

$$R_{DCE}^{\cdot\cdot\cdot A} = 2\delta_{[D}\Gamma^{A}{}_{C]E} + 2\Gamma^{A}_{[D|F|}\,\Gamma^{F}{}_{C]E}, \tag{4.4}$$

where $[DC]$; $[D \mid F \mid C]$ implies that the alternating product is to be formed on D, C.

We are now in a position to prove

THEOREM 3. *The metric tensor a_{CD} of S_m in K_n has the following properties*: (1) *the components a_{cd} (and $a_{c^*d^*}$) are analytic*[4] *functions of the surface parameters;* (2) *if the components of a_{cd} (and $a_{c^*d^*}$) are of rank m in some coördinate system, then they are related to the components $\Gamma^{e}{}_{fc}$ (and $\Gamma^{e^*}{}_{f^*c^*}$) in that coördinate system by the same relation as exists in Riemannian space*[5]*;* (3) *if the components a_{c^*d} (and a_{cd^*}) are of rank m in some coördinate system, then they are related to the components $\Gamma^{e}{}_{fc}$ (and $\Gamma^{e^*}{}_{f^*c^*}$) in that coördinate system by the same relation as the metric tensor of K_n is related to its connection.*[6]

By expanding (4.3), we obtain the relations

$$\begin{aligned}(\partial_f a_{cd})\,d\eta^f + (\partial_{f^*} a_{cd})\,d\eta^{f^*} - \Gamma^{e}{}_{fc} a_{ed}\,d\eta^f - \Gamma^{e}{}_{fd} a_{ce}\,d\eta^f = 0, \quad \text{conj.}\\ (\partial_f a_{c^*d})\,d\eta^f + (\partial_{f^*} a_{c^*d})\,d\eta^{f^*} - \Gamma^{e}{}_{fd} a_{c^*e}\,d\eta^f - \Gamma^{e^*}{}_{f^*c^*}\, a_{e^*d}\,d\eta^{f^*} = 0, \quad \text{conj.}\end{aligned} \tag{4.5}$$

Since the displacement vector $d\eta^E \rightarrow (d\eta^e,\, d\eta^{e^*})$ is arbitrary, we can assign to it: (1) the arbitrary real components $(\alpha, \beta, \gamma, \cdots, \alpha, \beta, \gamma, \cdots)$; (2) the arbitrary pure imaginary components $(i\alpha, i\beta, i\gamma, \cdots, -i\alpha, -i\beta, -i\gamma, \cdots)$. Upon replacing $(d\eta^e, d\eta^{e^*})$ by these two systems of components in each equation of (4.5), it will be found that each equation of (4.5) splits into two equations. The new system is

$$\begin{aligned}\partial_{f^*} a_{cd} = 0, \quad \text{conj.}; \qquad \partial_f a_{cd} - \Gamma^{e}{}_{fc} a_{ed} - \Gamma^{e}{}_{fd} a_{ce} = 0, \quad \text{conj.};\\ \partial_{f^*} a_{c^*d} - \Gamma^{e^*}{}_{f^*c^*} a_{e^*d} = 0, \quad \text{conj.}; \qquad \partial_f a_{c^*d} - \Gamma^{e}{}_{fd} a_{c^*e} = 0, \quad \text{conj.}\end{aligned} \tag{4.6}$$

The first equation of (4.6) leads to the first conclusion in our theorem. Since this relation is invariant under the group of S_m, it appears as if the group of S_m is the analytic group. In order to obtain the remaining conclusions, we assume that: (1) at least one coördinate system exists in which the rank of a_{cd} is m; (2) at least one coördinate system exists in which the rank of a_{c^*d} is m.

The coördinate systems of (1) and (2) need not coincide. Note that, in general, the ranks of a_{cd}, a_{c^*d} are not invariant under the group of S_m. In the specific coördinate systems of (1) and (2), we introduce the symmetric quantities b^{cd}, b^{cd^*} defined by

$$(4.7) \qquad b^{cd}a_{de} = {B^c}_e, \qquad b^{cd^*}a_{d^*e} = {B^c}_e, \quad \text{conj.}$$

Again, we stress the fact that (4.7) are not invariant relations. However, with the aid of these quantities, we find that in the coördinate systems of (1) and (2),

$$(4.8) \qquad {\Gamma^e}_{dc} = \tfrac{1}{2}b^{ef}[\partial_d a_{fc} + \partial_c a_{fd} - \partial_f a_{dc}], \quad \text{conj.},$$

$$(4.9) \qquad {\Gamma^e}_{dc} = b^{ef^*}\partial_d a_{f^*c}, \quad \text{conj.}$$

The relations (4.8), (4.9) are equivalent to the statements (2), (3) of our theorem.

As a direct consequence of equation (4.8), we obtain

THEOREM 4. *If one coördinate system exists in S_m for which the rank of a_{cd} is m, then in all coördinate systems ${\Gamma^e}_{dc}$ (and ${\Gamma^{e^*}}_{d^*c^*}$) are symmetric.*

If a coördinate system exists in S_m for which the rank of a_{cd} is m, then by (4.7) b^{cd} exists. Hence (4.8) is valid in that coördinate system. Since a_{cd} is symmetric in all coördinate systems, it follows from (4.8) that in our particular coördinate system ${\Gamma^e}_{dc}$ is symmetric. Now ${\Gamma^e}_{dc}$ (and ${\Gamma^{e^*}}_{d^*c^*}$) are the only non vanishing components of the connection of S_m. Hence the connection of S_m is symmetric in the above specified coördinate system. The components of the connection are transformed under the group of S_m by

$$(4.10) \qquad {\Gamma'^{C'}}_{B'D'} = B'^{CBD}_{C'B'D'}\,{\Gamma^C}_{BD} + {B'^{C'}}_C\,\partial_{B'}{B^C}_{D'}.$$

Evidently, if ${\Gamma^C}_{BD}$ is symmetric in one coördinate system, then it is symmetric in all coördinate systems. Thus, our theorem is established.

By use of (4.8), (4.9), we can prove

THEOREM 5. *If one coördinate system exists in S_m for which the ranks of both a_{cd}, a_{c^*d} are m, then the Riemann-Christoffel affinor of S_m vanishes.*

From our assumption, the equation (4.8) is valid. Since a_{cd} is analytic, it follows that b^{cd}, $\partial_f a_{cd}$ are analytic. Hence ${\Gamma^e}_{cd}$ is analytic, that is, in the above coördinate system,

$$(4.11) \qquad \partial_{f^*}{\Gamma^e}_{cd} = 0, \quad \text{conj.}$$

Furthermore, by our assumption, the equation (4.9) is valid. By direct differentiation of (4.9) and substitution into (4.4), it follows that

$$R^{\cdot\cdot\cdot a}_{dce} = 0, \quad \text{conj.} \tag{4.12}$$

Since (4.11) is equivalent to

$$R^{\cdot\cdot\cdot c}_{f^* ed} = 0, \quad \text{conj.}, \tag{4.13}$$

and all other components of $R^{\cdot\cdot\cdot A}_{DCE}$ vanish, it follows that in the above coördinate system

$$R^{\cdot\cdot\cdot A}_{DCE} = 0. \tag{4.14}$$

Since (4.14) is a tensor equation, it is valid in all coördinate systems of S_m. Note that it follows from this argument that (4.11) is a tensor equation. The invariance of (4.11) strongly suggests that the group of S_m is the analytic group.

From Theorems 4 and 5, we obtain

THEOREM 6. *If one coördinate system exists in S_m for which a_{cd}, a_{c^*d} are both of rank m, then S_m is a unitary Euclidean space.*

By definition,[7] a unitary Euclidean space is a unitary space for which Γ^C_{BD} is symmetric and $R^{\cdot\cdot\cdot A}_{DCE}$ vanishes.

UNIVERSITY OF TEXAS.

REFERENCES

Since all our references are to Schouten-Struik, *Einführung in die neueren Methoden der Differentialgeometrie*, vols. I, II (1935, 1938), we shall indicate merely volume and page.

[1] II, pp. 243-249.

[2] II, p. 234. For a complete introduction to the notation and fundamental concepts of Hermitian geometry, the reader is referred to this section of vol. II. We shall follow the notation of this section with two exceptions: (1) the star will replace the bar above the indices; (2) the star will be omitted from all letters unless ambiguity of meaning arises.

[3] II, p. 57.

[4] II, p. 226.

[5] I, p. 83.

[6] II, p. 234.

[7] II, p. 237.

BISYMMETRIC TENSOR ALGEBRA. PART I.*

By Richard H. Bruck and T. L. Wade.

Introduction. In general the literature of tensor algebra is relatively incomplete in a number of respects. For example, the concepts of determinant, adjoint, and inverse have received little attention. The object of the present paper is to develop these concepts in considerable detail for tensors of the form $A^{(i)}_{(j)}$, where $(i) = i_1 i_2 \cdots i_p$; that is, for tensors with p contravariant and p covariant indices with respect to an n-dimensional coördinate system. The components, or elements, of the tensors lie in some field, which we shall assume to be of characteristic zero, although parts of our theory will go through without this restriction. The algebra of all such tensors, for a given p, will be referred to as a total $2p$-tensor algebra, the tensors being of order $2p$.

A sub-algebra of the total $2p$-tensor algebra may be obtained by restricting attention to those tensors which are bisymmetric; that is, to tensors $A^{(i)}_{(j)} = A^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_p}$ which are left invariant when both the i's and the j's are subjected to the same arbitrary permutation. Because of the significant place of bisymmetric tensors in both pure and applied mathematics [1], [2], the authors were prompted to devote special attention to this sub-algebra. One advantage of a complete, self-contained treatment of the algebra of bisymmetric tensors is that it provides a natural and vivid realization of a semi-simple algebra.

We use in this paper, along with the principles of abstract algebra, the concepts of rank and rank tensor developed in Part I. With the aid of these concepts we are able in Part II to give an adequate theory of bisymmetric tensors in terms of the immanant tensors introduced by one of the writers in a previous paper [3].

1. Total $2p$-tensor algebra. Familiarity with the concept of a tensor, and with the elementary operations of addition of tensors and multiplication of a tensor by a scalar are assumed [4]. In view of the fact that we use multiplication of tensors in a special sense in this paper, we give particular attention to this operation.

The ordered product $C^{(i)}_{(j)}$ of two tensors $A^{(i)}_{(j)}$ and $B^{(i)}_{(j)}$ is defined by

$$A^{(i)}_{(m)} B^{(m)}_{(j)} = C^{(i)}_{(j)}. \tag{1}$$

This product is evidently unique and non-commutative.

---

* Received April 29, 1941; Presented to the American Mathematical Society, May 2, 1941, and abstracted in the *Bulletin of the Society*, vol. 47, 1941, pp. 373-374.

That an algebra of $2p$-tensors with the product defined in this manner might be of interest was indicated by the property of immanant tensors [3], which are mutually orthogonal idempotents in the sense that

$$ {}_{\alpha}I^{(i)}_{(m)}\;{}_{\beta}I^{(m)}_{(j)} = {}_{\alpha}I^{(i)}_{(j)} \quad \text{or} \quad 0 $$

according as $\alpha = \beta$ or not. Here, as elsewhere, we represent the null tensor by 0.

If $\delta_j{}^i$ is the simple Kronecker delta, it is evident that

$$ (2) \qquad \delta^{(i)}_{(j)} \equiv \delta^{i_1}_{j_1}\,\delta^{i_2}_{j_2}\cdots\delta^{i_p}_{j_p} $$

is the unit tensor under the multiplication defined by (1); for

$$ (3) \qquad A^{(i)}_{(m)}\,\delta^{(m)}_{(j)} = \delta^{(i)}_{(m)}\,A^{(m)}_{(j)} = A^{(i)}_{(j)}. $$

The $2p$-tensors, subject to the operations mentioned above, constitute a total linear associative algebra.

2. Extension of the use of indices. In order to define the determinant, adjoint, and inverse of a tensor $A^{(i)}_{(j)}$ we need to extend certain familiar ideas and principles. In this section the summation convention will be used but the entities considered will not be assumed to have tensor character.

Let N be an integer and let $\alpha_1, \alpha_2 \cdots \alpha_N$ be N distinct fixed labels. These need not be integers; in fact in the applications each label will consist of a set of integers. Let $R, S, T, R_1, S_1, T_1, \cdots$ be labels capable of ranging over the fixed labels $\alpha_1, \alpha_2, \cdots \alpha_N$. Finally let $A_T{}^S$ be a set of N^2 elements of the underlying field. We shall speak of $A_T{}^S$ as a matrix, and define the sum, products, etc. of two matrices in a manner which should be clear from the ordinary theory of matrices, if we substitute in the latter the range $\alpha_1, \alpha_2, \cdots \alpha_N$ for the range $1, 2, \cdots, N$. The unit matrix will be designated by "the simple Kronecker delta" $\delta_T{}^S$; that is

$$ (4) \qquad \delta_T{}^S = 0, 1 \qquad \text{according as } S \neq T, \quad S = T. $$

The determinant of a matrix $A_T{}^S$ will be defined by

$$ (5) \qquad |A| = |A_T{}^S| = \begin{vmatrix} A^{\alpha_1}_{\alpha_1} & A^{\alpha_1}_{\alpha_2} & \cdots & A^{\alpha_1}_{\alpha_N} \\ A^{\alpha_2}_{\alpha_1} & A^{\alpha_2}_{\alpha_2} & \cdots & A^{\alpha_2}_{\alpha_N} \\ \cdot & \cdot & \cdots & \cdot \\ A^{\alpha_N}_{\alpha_1} & A^{\alpha_N}_{\alpha_2} & \cdots & A^{\alpha_N}_{\alpha_N} \end{vmatrix}. $$

It is clear without comment that the usual product rule for determinants will hold; if

$$ (6) \qquad A_R{}^S\, B_T{}^R = C_T{}^S, $$

then

(7) $$|A| \cdot |B| = |C|.$$

As is usual in tensor algebra we may introduce the "generalized Kronecker delta"

(8) $$\delta^{S_1S_2\cdots S_N}_{T_1T_2\cdots T_N} = \begin{vmatrix} \delta^{S_1}_{T_1} & \delta^{S_1}_{T_2} & \cdots & \delta^{S_1}_{T_N} \\ \delta^{S_2}_{T_1} & \delta^{S_2}_{T_2} & \cdots & \delta^{S_2}_{T_N} \\ \cdot & \cdot & \cdots & \cdot \\ \delta^{S_N}_{T_1} & \delta^{S_N}_{T_2} & \cdots & \delta^{S_N}_{T_N} \end{vmatrix}.$$

In terms of this symbol it may be shown that

(9) $$|A_T{}^S| = \frac{1}{N!}\,\delta^{S_1S_2\cdots S_N}_{T_1T_2\cdots T_N}\,A^{T_1}_{S_1}\,A^{T_2}_{S_2}\cdots A^{T_N}_{S_N}.$$

Again, the adjoint $\bar{\mathfrak{A}}_T{}^S$ of a matrix $A_T{}^S$ may be defined by

(10) $$\bar{\mathfrak{A}}_T{}^S = \frac{1}{(N-1)!}\,\delta^{SS_2\cdots S_N}_{TT_2\cdots T_N}\,A^{T_2}_{S_2}\,A^{T_3}_{S_3}\cdots A^{T_N}_{S_N};$$

and the inverse of a non-singular matrix by

(11) $$\mathfrak{A}_T{}^S = \frac{1}{|A|}\cdot\bar{\mathfrak{A}}_T{}^S, \qquad |A| \neq 0.$$

We find it useful to introduce the concept of a rank matrix. If $k \overset{=}{<} N$ is any fixed integer, the set of quantities

(12) $$A^{S_1S_2\cdots S_k}_{T_1T_2\cdots T_k} = \begin{vmatrix} A^{S_1}_{T_1} & A^{S_1}_{T_2} & \cdots & A^{S_1}_{T_k} \\ A^{S_2}_{T_1} & A^{S_2}_{T_2} & \cdots & A^{S_2}_{T_k} \\ \cdot & \cdot & \cdots & \cdot \\ A^{S_k}_{T_1} & A^{S_k}_{T_2} & \cdots & A^{S_k}_{T_k} \end{vmatrix}$$

comprises the totality of all k by k minors of the determinant $|A_T{}^S|$; this set may also be regarded as a matrix, and we shall adopt this viewpoint. Making use of the usual definition of the rank of a two-way matrix or determinant, let us suppose that $A_T{}^S$ has rank r. Clearly then the matrix (12) vanishes unless $k \overset{=}{<} r$, while the matrix obtained for $k = r$ has at least one non-zero component.

Definition. If a matrix $A_T{}^S$ has rank r, then $A^{S_1S_2\cdots S_r}_{T_1T_2\cdots T_r}$ is the *rank matrix* of $A_T{}^S$.

THEOREM I. *Let the rank matrix* $A^{S_1S_2\cdots S_r}_{T_1T_2\cdots T_r}$ *of a matrix* $A_T{}^S$ *have the non-zero component*

(13) $$c = A^{\beta_1\beta_2\cdots\beta_r}_{\gamma_1\gamma_2\cdots\gamma_r} \neq 0,$$

the β's and γ's being fixed labels. Then

$$A^{S_1S_2\cdots S_r}_{T_1T_2\cdots T_r} = \frac{1}{c}\cdot A^{S_1S_2\cdots S_r}\cdot A_{T_1T_2\cdots T_r}, \tag{14}$$

where we define

$$A^{S_1S_2\cdots S_r} = A^{S_1S_2\cdots S_r}_{\gamma_1\gamma_2\cdots\gamma_r}; \quad A_{T_1T_2\cdots T_r} = A^{\beta_1\beta_2\cdots\beta_r}_{T_1T_2\cdots T_r}. \tag{15}$$

Proof. Since $A_T{}^S$ has rank r, the determinant (5) has exactly r linearly independent columns. According to (13), the columns labelled $\gamma_1, \gamma_2, \cdots, \gamma_r$ form a linearly independent set. Hence

$$A_T{}^S = A_{\gamma_1}{}^S u_T{}^{\gamma_1} + A_{\gamma_2}{}^S u_T{}^{\gamma_2} + \cdots + A_{\gamma_r}{}^S u_T{}^{\gamma_r}, \tag{16}$$

where the u's are a set of elements which it will not be necessary to evaluate. Because of (16) it is clear by the rule for the multiplication of determinants that

$$A^{S_1S_2\cdots S_r}_{T_1T_2\cdots T_r} = A^{S_1S_2\cdots S_r}\cdot \begin{vmatrix} u^{\gamma_1}_{T_1} & u^{\gamma_1}_{T_2} & \cdots & u^{\gamma_1}_{T_r} \\ \cdot & \cdot & \cdots & \cdot \\ u^{\gamma_r}_{T_1} & u^{\gamma_r}_{T_2} & \cdots & u^{\gamma_r}_{T_r} \end{vmatrix} \tag{17}$$

in which $A^{S_1S_2\cdots S_r}$ has been defined by (15). Now let us put $S_1S_2\cdots S_r = \beta_1\beta_2\cdots\beta_r$; from (17), (15), and (13) we obtain

$$A_{T_1T_2\cdots T_r} = c\cdot \begin{vmatrix} u^{\gamma_1}_{T_1} & \cdots & u^{\gamma_1}_{T_r} \\ \cdot & \cdots & \cdot \\ u^{\gamma_r}_{T_1} & \cdots & u^{\gamma_r}_{T_r} \end{vmatrix}. \tag{18}$$

Then (18) and (17) combined give (14).

COROLLARY. *The rank of the rank matrix of a non-zero matrix is unity.*

Proof. If we use the notation $(S) = S_1S_2\cdots S_r$, then the rank matrix may be denoted by $A^{(S)}_{(T)}$, and equation (14) becomes $A^{(S)}_{(T)} = \frac{1}{c}\cdot A^{(S)}\cdot A_{(T)}$. Once we have chosen a proper set of fixed labels for the range of (S), we may define a determinant of $A^{(S)}_{(T)}$ and a corresponding rank. That this rank is unity follows from the fact that

$$\begin{vmatrix} A^{(S_1)}_{(T_1)} & A^{(S_1)}_{(T_2)} \\ A^{(S_2)}_{(T_1)} & A^{(S_2)}_{(T_2)} \end{vmatrix} = \frac{1}{c^2}\cdot A^{(S_1)}\cdot A^{(S_2)}\cdot A_{(T_1)}\cdot A_{(T_2)}\cdot \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0.$$

When S, T take on the range $1, 2, \cdots, N$ instead of the range $\alpha_1, \alpha_2, \cdots, \alpha_N$ this theorem becomes Kowalewski's theorem [5].

As a simple example of the application of the theorem, we note that $\delta_T{}^S$ has determinant 1 and hence has rank N. Thus the "generalized Kronecker delta" is the rank matrix of $\delta_T{}^S$, and we have

$$\delta^{S_1S_2\cdots S_N}_{T_1T_2\cdots T_N} = \delta^{S_1S_2\cdots S_N}\cdot\delta_{T_1T_2\cdots T_N}; \tag{19}$$

here

$$\begin{cases} \delta^{S_1S_2\cdots S_N} = \delta^{S_1S_2\cdots S_N}_{a_1a_2\cdots a_N} \\ \delta_{T_1T_2\cdots T_N} = \delta^{a_1a_2\cdots a_N}_{T_1T_2\cdots T_N}. \end{cases} \tag{20}$$

3. The determinant of $A^{(i)}_{(j)}$. The essential step is to identify the label S of Section 2 with the ordered set of tensor indices (i):

$$S = (i) = i_1i_2\cdots i_p. \tag{21}$$

As the p indices i run independently over the range 1 to n, S will take on $N = n^p$ fixed labels, which, when arranged in any assigned order, we may take to be the $\alpha_1, \alpha_2, \cdots, \alpha_N$ of the previous section.

Here $\delta_T{}^S = \delta^{(i)}_{(j)}$ is given by (2), Section 1; and hence, from (8),

$$\delta^{(i_1)(i_2)\cdots(i_N)}_{(j_1)(j_2)\cdots(j_N)} = \begin{vmatrix} \delta^{(i_1)}_{(j_1)} & \delta^{(i_1)}_{(j_2)} & \cdots & \delta^{(i_1)}_{(j_N)} \\ \delta^{(i_2)}_{(j_1)} & \delta^{(i_2)}_{(j_2)} & \cdots & \delta^{(i_2)}_{(j_N)} \\ \cdot & \cdot & \cdots & \cdot \\ \delta^{(i_N)}_{(j_1)} & \delta^{(i_N)}_{(j_2)} & \cdots & \delta^{(i_N)}_{(j_N)} \end{vmatrix}. \tag{22}$$

We use

$$(i_\lambda) = i_{\lambda_1}i_{\lambda_2}\cdots i_{\lambda_p};$$

this notation being a simple extension of that of the Introduction. When $p = 1$, so that (i_λ) may be replaced by i_λ, and $N = n^p = n$, (22) becomes

$$\delta^{i_1i_2\cdots i_n}_{j_1j_2\cdots j_n} = \begin{vmatrix} \delta^{i_1}_{j_1} & \delta^{i_1}_{j_2} & \cdots & \delta^{i_1}_{j_n} \\ \delta^{i_2}_{j_1} & \delta^{i_2}_{j_2} & \cdots & \delta^{i_2}_{j_n} \\ \cdot & \cdot & \cdots & \cdot \\ \delta^{i_n}_{j_1} & \delta^{i_n}_{j_2} & \cdots & \delta^{i_n}_{j_n} \end{vmatrix}. \tag{23}$$

This last, studied by Murnaghan [6] and others [7] is known as the generalized Kronecker delta. Evidently the tensor defined by (22) constitutes a generalization of that defined by (23); we propose to extend the term *generalized Kronecker delta* to apply to (22) as well.

The generalized Kronecker delta defined by (22) has the following values:

(1) 0, if two or more of the N sets of subscripts (or superscripts) are equal.

(2) 1, if the N sets of subscripts differ from the N sets of superscripts by an even permutation;

(3) -1, if the N sets of subscripts differ from the N sets of superscripts by an odd permutation.

Equivalently we might say that the generalized Kronecker delta is skew-symmetric in the N sets of superscripts and in the N sets of subscripts, and that $\delta^{a_1 a_2 \cdots a_N}_{a_1 a_2 \cdots a_N} = 1$.

If $k \leqq N$, we may define

$$(24) \qquad \delta^{(i_1)\cdots(i_k)}_{(j_1)\cdots(j_k)} = \begin{vmatrix} \delta^{(i_1)}_{(j_1)} & \cdots & \delta^{(i_1)}_{(j_k)} \\ \cdot & \cdots & \cdot \\ \delta^{(i_k)}_{(j_1)} & \cdots & \delta^{(i_k)}_{(j_k)} \end{vmatrix}.$$

The identity

$$(25) \qquad \delta^{(i_1)(i_2)\cdots(i_k)(i_{k+1})\cdots(i_N)}_{(j_1)(j_2)\cdots(j_k)(i_{k+1})\cdots(i_N)} = (N-k)!\,\delta^{(i_1)(i_2)\cdots(i_k)}_{(j_1)(j_2)\cdots(j_k)}$$

readily follows. In particular we single out the indentities

$$(26) \qquad \delta^{(i_1)(i_2)\cdots(i_N)}_{(j_1)(i_2)\cdots(i_N)} = (N-1)!\,\delta^{(i_1)}_{(j_1)}$$

and

$$(27) \qquad \delta^{(i_1)(i_2)\cdots(i_N)}_{(i_1)(i_2)\cdots(i_N)} = N!\,.$$

For a tensor $A^{(i)}_{(j)}$ and any particular integer k we may define a tensor $A^{(i_1)\cdots(i_k)}_{(j_1)\cdots(j_k)}$ in the manner indicated by (12). If r is the greatest value of k for which the latter tensor does not vanish, we shall call r the *rank* of $A^{(i)}_{(j)}$ and $A^{(i_1)\cdots(i_k)}_{(j_1)\cdots(j_k)}$ the *rank tensor* of $A^{(i)}_{(j)}$. Evidently r is an arithmetic invariant. For a given p the unit tensor for the corresponding total $2p$-tensor algebra is $\delta^{(i)}_{(j)}$, given by (2); and the rank tensor of $\delta^{(i)}_{(j)}$ is the generalized Kronecker delta $\delta^{(i_1)\cdots(i_N)}_{(j_1)\cdots(j_N)}$ for this algebra.

With (19), (20), and (21) in mind we have

$$(28) \qquad \delta^{(i_1)(i_2)\cdots(i_N)}_{(j_1)(j_2)\cdots(j_N)} = \delta^{(i_1)(i_2)\cdots(i_N)} \cdot \delta_{(j_1)(j_2)\cdots(j_N)}$$

where

$$(29) \qquad \begin{cases} \delta^{(i_1)(i_2)\cdots(i_N)} = \delta^{(i_1)(i_2)\cdots(i_N)}_{a_1\ a_2\ \cdots\ a_N} \\ \delta_{(j_1)(j_2)\cdots(j_N)} = \delta^{a_1\ a_2\ \cdots\ a_N}_{(j_1)(j_2)\cdots(j_N)}. \end{cases}$$

The two entities defined by (29) are skew-symmetric in the N sets of indices; moreover $\delta^{a_1 a_2 \cdots a_N} = \delta_{a_1 a_2 \cdots a_N} = 1$.

Although $\delta^{(i_1)\cdots(i_N)}_{(j_1)\cdots(j_N)}$, defined by (22) is unmistakably an absolute tensor, since it is an algebraic sum of absolute tensors of the same order, it is less evident that the entities defined by (29) are also tensors. We shall prove the latter point presently.

Definition. We define the *determinant* $|A| = |A^{(i)}_{(j)}|$ of $A^{(i)}_{(j)}$ by

$$(30)\qquad |A| = \frac{1}{N!}\,\delta^{(i_1)(i_2)\ldots(i_N)}_{(j_1)(j_2)\ldots(j_N)}\,A^{(j_1)}_{(i_1)}\,A^{(j_2)}_{(i_2)}\cdots A^{(j_N)}_{(i_N)}.$$

LEMMA. *The determinant, $|A|$, satisfies the relations*

$$(31)\qquad \delta_{(r_1)(r_2)\ldots(r_N)}\cdot|A| = \delta_{(j_1)(j_2)\ldots(j_N)}\,A^{(j_1)}_{(r_1)}\,A^{(j_2)}_{(r_2)}\cdots A^{(j_N)}_{(r_N)};$$

$$(31a)\qquad \delta^{(r_1)(r_2)\ldots(r_N)}\cdot|A| = \delta^{(j_1)(j_2)\ldots(j_N)}\,A^{(r_1)}_{(j_1)}\,A^{(r_2)}_{(j_2)}\cdots A^{(r_N)}_{(j_N)}.$$

Proof. We give the proof for (31). Multiplying both sides of (30) by $\delta_{(r_1)\ldots(r_N)}$, and using (28), we obtain

$$\delta_{(r_1)\ldots(r_N)}\cdot|A| = \frac{1}{N!}\,\delta^{(i_1)\ldots(i_N)}_{(r_1)\ldots(r_N)}\cdot\left[\delta_{(j_1)\ldots(j_N)}\,A^{(j_1)}_{(i_1)}\cdots A^{(j_N)}_{(i_N)}\right].$$

Since the entity in square brackets is skew-symmetric in the (i)'s in the same way that $\delta^{(i_1)\ldots(i_N)}_{(r_1)\ldots(r_N)}$ is skew-symmetric in the (r)'s, (31) follows.

THEOREM II. $\delta_{(i_1)\ldots(i_N)}$ *and* $\delta^{(i_1)\ldots(i_N)}$ *are relative tensors of weights* w *and* $-w$ *respectively, where* $w = pN/n$.

Proof. In (31) set

$$A^{(j)}_{(r)} = t^{j_1}_{r_1}\,t^{j_2}_{r_2}\cdots t^{j_p}_{r_p}.$$

Thus if

$$\delta'_{(r_1)\ldots(r_N)} = \delta_{(j_1)\ldots(j_N)}\cdot\left[t^{j_{11}}_{r_{11}}\,t^{j_{12}}_{r_{12}}\cdots t^{j_{1p}}_{r_{1p}}\right]\cdots\left[t^{j_{N1}}_{r_{N1}}\,t^{j_{N2}}_{r_{N2}}\cdots t^{j_{Np}}_{r_{Np}}\right]$$

(31) states that

$$\delta'_{(r_1)\ldots(r_N)} = \left|\,t^{i_1}_{j_1}\,t^{i_2}_{j_2}\cdots t^{i_p}_{j_p}\right|\cdot\delta_{(r_1)\ldots(r_N)}.$$

Thus $\delta_{(r_1)\ldots(r_N)}$ transforms as a relative tensor, and a simple caculation shows that its weight $w = pN/n$. The proof for $\delta^{(i_1)\ldots(i_N)}$ follows similar lines.

It is implicit in Cramlet's work [8] that any relative tensor such as $\delta^{(i_1)\ldots(i_N)}$ is a rational integral function of the relative tensor $\delta^{i_1 i_2\cdots i_n} = \epsilon^{i_1 i_2\cdots i_n}$. As an illustration, let $n = p = 2$; $N = 2^2 = 4$; here the unit tensor is $\delta^{(i)}_{(j)} = \delta^{i_1}_{j_1}\,\delta^{i_2}_{j_2}$. A fourth order determinant expression for $\epsilon^{(i_1)(i_2)(i_3)(i_4)}$ is obtained from (22) by putting $(j_1) = 1, 1$; $(j_2) = 1, 2$; $(j_3) = 2, 1$; $(j_4) = 2, 2$. Expanding this determinant by two-rowed minors in the direction of the columns, and using the identity $\epsilon^{ab}\epsilon^{cd} = \epsilon^{ac}\epsilon^{bd} + \epsilon^{cb}\epsilon^{ad}$, we obtain

$$\delta^{(i_1)(i_2)(i_3)(i_4)} = \begin{vmatrix} \epsilon^{i_{12}i_{22}}\cdot\epsilon^{i_{32}i_{42}}, & \epsilon^{i_{11}i_{31}}\cdot\epsilon^{i_{21}i_{41}} \\ \epsilon^{i_{12}i_{32}}\cdot\epsilon^{i_{22}i_{42}}, & \epsilon^{i_{11}i_{21}}\cdot\epsilon^{i_{31}i_{41}} \end{vmatrix}.$$

THEOREM III. *The determinant of the product of two tensors is equal to the product of their determinants.*

Proof. Let

$$C^{(i)}_{(j)} = A^{(i)}_{(m)} B^{(m)}_{(j)}.$$

Then

$$\begin{aligned}
|C| &= \frac{1}{N!}\delta^{(i_1)\cdots(i_N)}_{(j_1)\cdots(j_N)} C^{(j_1)}_{(i_1)}\cdots C^{(j_N)}_{(i_N)}\\
&= \frac{1}{N!}\delta^{(i_1)\cdots(i_N)}_{(j_1)\cdots(j_N)} A^{(j_1)}_{(m_1)} B^{(m_1)}_{(i_1)}\cdots A^{(j_N)}_{(m_N)} B^{(m_N)}_{(i_N)}\\
&= \frac{1}{N!}\delta^{(i_1)\cdots(i_N)}\cdot\left[\delta_{(j_1)\cdots(j_N)}\cdot A^{(j_1)}_{(m_1)}\cdots A^{(j_N)}_{(m_N)}\right]\cdot B^{(m_1)}_{(i_1)}\cdots B^{(m_N)}_{(i_N)}\\
&= \frac{1}{N!}\delta^{(i_1)\cdots(i_N)}\cdot\delta_{(m_1)\cdots(m_N)}\cdot|A|\cdot B^{(m_1)}_{(i_1)}\cdots B^{(m_N)}_{(i_N)} \qquad \text{(by (31))}\\
&= |A|\cdot\frac{1}{N!}\delta^{(i_1)\cdots(i_N)}_{(m_1)\cdots(m_N)} B^{(m_1)}_{(i_1)}\cdots B^{(m_N)}_{(i_N)}\\
&= |A|\cdot|B|. \qquad \text{(by (30))}
\end{aligned}$$

or

$$|C| = |A|\cdot|B|.$$

4. The adjoint and inverse of $A^{(i)}_{(j)}$. The cofactor $\bar{\mathfrak{A}}^{(i_1)}_{(j_1)}$ of $A^{(j_1)}_{(i_1)}$ in $|A|$ is given by

$$\bar{\mathfrak{A}}^{(i_1)}_{(j_1)} = \frac{1}{(N-1)!}\delta^{(i_1)(i_2)\cdots(i_N)}_{(j_1)(j_2)\cdots(j_N)} A^{(j_2)}_{(i_2)}\cdots A^{(j_N)}_{(i_N)}. \tag{32}$$

Forming the product of this with $A^{(j_1)}_{(i_1)}$, we get

$$\begin{aligned}
\bar{\mathfrak{A}}^{(i_1)}_{(m_1)} A^{(m_1)}_{(j_1)} &= \frac{1}{(N-1)!}\delta^{(i_1)(i_2)\cdots(i_N)}_{(m_1)(j_2)\cdots(j_N)} A^{(m_1)}_{(j_1)} A^{(j_2)}_{(i_2)}\cdots A^{(j_N)}_{(i_N)}\\
&= \frac{1}{(N-1)!}\delta^{(i_1)(i_2)\cdots(i_N)}\cdot\left[\delta_{(m_1)(j_2)\cdots(j_N)} A^{(m_1)}_{(j_1)} A^{(j_2)}_{(i_2)}\cdots A^{(j_N)}_{(i_N)}\right]\\
&= \frac{1}{(N-1)!}\delta^{(i_1)(i_2)\cdots(i_N)}\cdot\delta_{(j_1)(i_2)\cdots(i_N)}\cdot|A| \qquad \text{(by (31))}\\
&= \frac{1}{(N-1)!}\delta^{(i_1)(i_2)\cdots(i_N)}_{(j_1)(i_2)\cdots(i_N)}\cdot|A|\\
&= \frac{1}{(N-1)!}\cdot(N-1)!\,\delta^{(i_1)}_{(j_1)}\cdot|A|, \qquad \text{(by (26))}
\end{aligned}$$

or

$$\bar{\mathfrak{A}}^{(i)}_{(m)} A^{(m)}_{(j)} = |A|\cdot\delta^{(i)}_{(j)}. \tag{33}$$

By a similar procedure, using (31a), it may be established that

$$A^{(i)}_{(m)}\bar{\mathfrak{A}}^{(m)}_{(j)} = |A|\cdot\delta^{(i)}_{(j)}. \tag{33a}$$

The tensor $\bar{\mathfrak{A}}^{(i)}_{(j)}$, defined by (32), we call the *adjoint* of $A^{(i)}_{(j)}$.

The *inverse* of a non-singular tensor is defined by

$$\mathfrak{A}^{(i)}_{(j)} = \frac{1}{|A|}\cdot\bar{\mathfrak{A}}^{(i)}_{(j)}, \qquad (|A| \neq 0). \tag{34}$$

Obviously from (33) and (33a)

$$\mathfrak{A}^{(i)}_{(m)} A^{(m)}_{(j)} = A^{(i)}_{(m)} \mathfrak{A}^{(m)}_{(j)} = \delta^{(i)}_{(j)}. \tag{35}$$

5. Remarks. If the unit tensor $\delta^{(i)}_{(j)}$ be replaced by an idempotent numerical tensor $C^{(i)}_{(j)}$, and the total $2p$-tensor algebra by the subalgebra consisting of all tensors of the form $C^{(i)}_{(l)} A^{(l)}_{(m)} C^{(m)}_{(j)}$, it is natural to inquire concerning the formation of an inverse and determinant. One would hope that the rank tensor of $C^{(i)}_{(j)}$ would play the rôle of the generalized Kronecker delta $\delta^{(i_1)\cdots(i_N)}_{(j_1)\cdots(j_N)}$ as in formulas (30) and (34), and that there would be an analogue of formula (25).

In Part II these hopes and anticipations are fully realized upon the groundwork of the present paper.

UNIVERSITY OF ALABAMA.

REFERENCES

[1] H. Weyl, *The Classical Groups*, Princeton, 1939. Especially p. 98.

[2] ———, *The Theory of Groups and Quantum Mechanics*, New York, 1932. Chapter V.

[3] T. L. Wade, "Tensor algebra and Young's symmetry operators," *American Journal of Mathematics*, vol. 63 (1941), pp. 645-657.

[4] H. Rothe, *Einfuhrung in die Tensorrechnung*, Wien, 1924.

[5] G. Kowalewski, *Einfuhrung in die Determinantentheorie*, Berlin, 1925, p. 113.

[6] F. D. Murnaghan, "The tensor character of the generalized Kronecker symbol," *Bulletin of the American Mathematical Society*, vol. 31 (1925), pp. 323-334.

[7] O. Veblen, *Invariants of Quadratic Differential Forms, Cambridge Tracts*, no. 24, p. 12.

[8] C. M. Cramlet, "The derivation of algebraic invariants by tensor algebra," *Bulletin of the American Mathematical Society*, vol. 34 (1928), pp. 334-342.

BISYMMETRIC TENSOR ALGEBRA. PART II.[*]

By RICHARD H. BRUCK and T. L. WADE.

1. The group ring of the symmetric group on p letters and the algebra of absolute numerical tensors.

Let $\pi = \binom{1\ 2\ \cdots\ p}{\pi_1 \pi_2 \cdots \pi_p}$ and $\tau = \binom{1\ 2\ \cdots\ p}{\tau_1 \tau_2 \cdots \tau_p}$ be any two permutations of the symmetric group on p letters; we denote these briefly by $\pi = \binom{t}{\pi_t}$, $\tau = \binom{t}{\tau_t}$. Then $\pi\tau = \binom{t}{\tau_{\pi_t}}$, by the customary method of forming permutations [1]; the opposite order of multiplication is sometimes used (see Weyl [2]).

In this paper we use the terminology and results of Part I[1] and assume, as there, that the underlying field has characteristic zero.

From the unit tensor $\delta^{(i)}_{(j)} = \delta^{i_1}_{j_1} \delta^{i_2}_{j_2} \cdots \delta^{i_p}_{j_p}$ may be formed conjugate tensors, or isomers, of the type $T^{(i)}_{(j)} = \delta^{(i)}_{(j_\pi)} = \delta^{i_1}_{j_{\pi_1}} \delta^{i_2}_{j_{\pi_2}} \cdots \delta^{i_p}_{j_{\pi_p}}$. The product of $\delta^{(i)}_{(j_\pi)}$ and $\delta^{(i)}_{(j_\tau)}$, under the multiplication defined in Part I, is seen to be $\delta^{(i)}_{(j_{\pi\tau})}$:

$$\delta^{(i)}_{(m_\pi)} \delta^{(m)}_{(j_\tau)} = \delta^{(i)}_{(j_{\pi\tau})}. \tag{1}$$

Thus the set $\mathfrak{L}$ of all linear sums of the unit tensor and its isomers forms a tensor algebra. We now state two basic theorems.

THEOREM I. *Any absolute numerical tensor is, for some p, a member of the algebra $\mathfrak{L}$.*

This is merely a restatement of the results of Cramlet [3] and Knebelman [4].

THEOREM II. *For a given p the algebra $\mathfrak{L}$ comprises the commutator algebra of the algebra of all bisymmetric tensors, and conversely.*

Proof. If a tensor commutes with all bisymmetric tensors, it must, in particular, commute with all tensors of the form $A^{(i)}_{(j)} = t^{i_1}_{j_1} t^{i_2}_{j_2} \cdots t^{i_p}_{j_p}$; equivalently it must be an absolute numerical tensor and belong to $\mathfrak{L}$. Again, if $A^{(i)}_{(j)}$ is any tensor, not necessarily bisymmetric, and if $\pi' = \binom{\pi_1 \pi_2 \cdots \pi_p}{1\ 2\ \cdots\ p} = \binom{1\ 2\ \cdots\ p}{\pi'_1 \pi'_2 \cdots \pi'_p}$ is the inverse permutation of π, it follows that

$$A^{(i)}_{(m)} \delta^{(m)}_{(j_\pi)} = A^{(i)}_{(j_\pi)},$$

---

[*] Received April 20, 1941; Presented to the American Mathematical Society, May 2, 1941, and abstracted in the *Bulletin of the Society*, vol. 47 (1941), pp. 373-374.

[1] Richard H. Bruck and T. L. Wade, "Bisymmetric tensor algebra. Part I"; this JOURNAL, vol. 64 (1942), pp. 725-733.

This paper will be referred to as Part I.

but

$$\delta^{(i)}_{(m_\pi)}\, A^{(m)}_{(j)} = \delta^{(i_{\pi'})}_{(m)}\, A^{(m)}_{(j)} = A^{(i_{\pi'})}_{(i)}.$$

If $A^{(i)}_{(j)}$ is bisymmetric, $A^{(i_{\pi'})}_{(j)} = A^{(i)}_{(j_\pi)}$, and so $\mathfrak{L}$ commutes with all bisymmetric tensors. Conversely, if $A^{(i)}_{(j)}$ is to commute with $\mathfrak{L}$, the equation $A^{(i_{\pi'})}_{(j)} = A^{(i)}_{(j_\pi)}$, or the equivalent equation $A^{(i)}_{(j)} = A^{(i_\pi)}_{(j_\pi)}$, must hold for all π; that is $A^{(i)}_{(j)}$ must be bisymmetric.

It is important for our purpose to note the homomorphic correspondence between the group ring $\mathfrak{S}_p$ of the symmetric group on p letters and the algebra $\mathfrak{L}$. In fact we may set up the correspondence

$$a = \Sigma a(\pi) \cdot \pi \rightarrow \Sigma a(\pi) \cdot \delta^{(i)}_{(j_\pi)}, \tag{2}$$

where a is any member of $\mathfrak{S}_p$. In view of (1) the correspondence is a homomorphism. If n, the dimension of the coördinate system, is $\geqq p$, the correspondence is indeed an isomorphism, but when n is less than p this is not true. In fact, let $\lambda(\pi) = -1, +1$ according as the permutation π is odd or even. then under (2)

$$\Sigma\lambda(\pi) \cdot \pi \rightarrow \delta^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_p},$$

but the tensor on the right is zero for $n < p$. (See [5]).

2. The subalgebra defined by an idempotent. The theory of this section is a generalization of Section 3, Part I, the essential change being that the unit tensor $\delta^{(i)}_{(j)}$ is here replaced by a more general idempotent $C^{(i)}_{(j)}$. In so far as possible, the treatment of this section runs parallel to that of the former section. It has been necessary in some places to introduce new proofs. Often, however, a proof used in Part I may be applied to the present needs merely by replacing the letter δ by the letter C; in such a case we omit it in the interest of brevity.

Let $C^{(i)}_{(j)}$ be a numerical idempotent tensor; thus $C^{(i)}_{(j)}$ is an absolute numerical tensor satisfying the relation

$$C^{(i)}_{(m)} \cdot C^{(m)}_{(j)} = C^{(i)}_{(j)}. \tag{3}$$

Let r be the rank of $C^{(i)}_{(j)}$; then its rank tensor is

$$C^{(i_1)(i_2)\cdots(i_r)}_{(j_1)(j_2)\cdots(j_r)} = \begin{vmatrix} C^{(i_1)}_{(j_1)} & C^{(i_1)}_{(j_2)} & \cdots & C^{(i_1)}_{(j_r)} \\ C^{(i_2)}_{(j_1)} & C^{(i_2)}_{(j_2)} & \cdots & C^{(i_2)}_{(j_r)} \\ \cdot & \cdot & \cdots & \cdot \\ C^{(i_r)}_{(j_1)} & C^{(i_r)}_{(j_2)} & \cdots & C^{(i_r)}_{(j_r)} \end{vmatrix}. \tag{4}$$

As in Theorem I of Part I, this factors:

$$C^{(i_1)(i_2)\cdots(i_r)}_{(j_1)(j_2)\cdots(j_r)} = \frac{1}{c} \cdot C^{(i_1)(i_2)\cdots(i_r)} \cdot C_{(j_1)(j_2)\cdots(j_r)} \qquad (c \neq 0). \tag{5}$$

We should point out that we may have a certain freedom of choice in regard to the entities on the right hand side of (5). In fact, let $C^{\beta_1\beta_2\cdots\beta_r}_{\gamma_1\gamma_2\cdots\gamma_r}$ be a non-zero component of the rank-tensor; clearly there will be more than one choice of the superscript labels $\beta_1\beta_2\cdots\beta_r$ and of the subscript labels $\gamma_1\gamma_2\cdots\gamma_r$. Once our choice has been made

$$(5a)\qquad c = C^{\beta_1\cdots\beta_r}_{\gamma_1\cdots\gamma_r} \neq 0; \quad C^{(i_1)\cdots(i_r)} = C^{(i_1)\cdots(i_r)}_{\gamma_1\cdots\gamma_r}; \quad C_{(j_1)\cdots(j_r)} = C^{\beta_1\cdots\beta_r}_{(j_1)\cdots(j_r)}.$$

That the entities $C^{(i_1)\cdots(i_r)}$ and $C_{(j_1)\cdots(j_r)}$ are tensors will appear in the course of the work.

It seems desirable to make the following remarks of an anticipatory nature. We shall give particular attention to the algebra of bisymmetric tensors, which in the future we shall denote by $\mathfrak{B}$. For $A^{(i)}_{(j)} \subset \mathfrak{B}$ the set $\mathfrak{S}$ of all tensors $A^{(i)}_{(m)} C^{(m)}_{(j)} = C^{(i)}_{(m)} A^{(m)}_{(j)}$ is clearly a subalgebra of the total $2p$-tensor algebra with unit tensor $C^{(i)}_{(j)}$ but it is not in general a subalgebra of $\mathfrak{B}$. We shall find that when $C^{(i)}_{(j)}$ is an immanant tensor (and so lies in the centrum of the commutator algebra $\mathfrak{L}$) the corresponding algebra $\mathfrak{S}$ is, however, a subalgebra of $\mathfrak{B}$. For any $A^{(i)}_{(j)}, B^{(i)}_{(j)} \subset \mathfrak{B}$ the relations

$$B^{(i)}_{(l)} \cdot A^{(l)}_{(m)} C^{(m)}_{(j)} = B^{(i)}_{(l)} A^{(l)}_{(m)} \cdot C^{(m)}_{(j)}, \text{ and } A^{(i)}_{(m)} C^{(m)}_{(l)} \cdot B^{(l)}_{(j)} = A^{(i)}_{(m)} B^{(m)}_{(l)} \cdot C^{(l)}_{(j)}$$

hold; from this we may conclude that in case $\mathfrak{S}$ is a subalgebra of $\mathfrak{B}$ it is in addition an invariant subalgebra of $\mathfrak{B}$.

We define the *determinant* $\Delta(A) = \Delta_C(A)$ of a bisymmetric tensor $A^{(i)}_{(j)}$ *with respect to* $\mathfrak{S} = C\mathfrak{B}C$ by

$$(6)\qquad \Delta_C(A) = \frac{1}{r!} \cdot C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} \cdot A^{(j_1)}_{(i_1)} \cdots A^{(j_r)}_{(i_r)}.$$

LEMMA I. *Let* $T^{(i_1)(i_2)\cdots(i_r)}_{(j_1)(j_2)\cdots(j_r)}$ *be a tensor, skew-symmetric in the* r *sets of subscripts and in the* r *sets of superscripts, with the further properties:*

$$(a)\quad T^{(i_1)\cdots(i_r)}_{(m)\cdots(j_r)} C^{(m)}_{(j_1)} = T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}; \quad (b)\quad C^{(i_1)}_{(m)} T^{(m)\cdots(i_r)}_{(j_1)\cdots(j_r)} = T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}.$$

Then

$$(7)\qquad T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = \lambda C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$$

for some scalar λ.

Proof. If we consider the determinant (4) to be expanded as an algebraic sum of $r!$ terms, we may deduce from (a) and the skew-symmetry properties of $T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$ that

$$(8)\qquad T^{(i_1)\cdots(i_r)}_{(m_1)\cdots(m_r)} C^{(m_1)\cdots(m_r)}_{(j_1)\cdots(j_r)} = r!\, T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}.$$

Similarly, from (b),

$$(9)\qquad C^{(i_1)\cdots(i_r)}_{(l_1)\cdots(l_r)} T^{(l_1)\cdots(l_r)}_{(j_1)\cdots(j_r)} = r!\, T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}.$$

By combining (8) and (9) we have

$$T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = \frac{1}{(r!)^2} C^{(i_1)\cdots(i_r)}_{(l_1)\cdots(l_r)} T^{(l_1)\cdots(l_r)}_{(m_1)\cdots(m_r)} C^{(m_1)\cdots(m_r)}_{(j_1)\cdots(j_r)}.$$

Because of (5) the right hand side agrees with (7), where

$$\lambda = \frac{1}{(r!)^2} \cdot C^{(m_1)\cdots(m_r)}_{(l_1)\cdots(l_r)} T^{(l_1)\cdots(l_r)}_{(m_1)\cdots(m_r)}.$$

In other words, the properties ascribed to $T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$, and which are clearly possessed by $C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$, suffice to characterize the latter to within a scalar factor.

From (7) and (8) follows

$$C^{(i_1)\cdots(i_r)}_{(m_1)\cdots(m_r)} C^{(m_1)\cdots(m_r)}_{(j_1)\cdots(j_r)} = r!\, C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}.$$

By (5) this may be written

$$C^{(m_1)\cdots(m_r)}_{(m_1)\cdots(m_r)} \cdot C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = r!\, C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)};$$

whence

$$C^{(i_1)\cdots(i_r)}_{(i_1)\cdots(i_r)} = r!. \tag{10}$$

Lᴇᴍᴍᴀ II. *The determinant* $\Delta_C(A)$, *defined by* (6), *where* $A^{(i)}_{(j)}$ *is bisymmetric, satisfies the relations*:

$$C_{(j_1)\cdots(j_r)} \cdot \Delta_C(A) = C_{(m_1)\cdots(m_r)} A^{(m_1)}_{(j_1)} \cdots A^{(m_r)}_{(j_r)}; \tag{11}$$

$$C^{(i_1)\cdots(i_r)} \cdot \Delta_C(A) = C^{(m_1)\cdots(m_r)} A^{(i_1)}_{(m_1)} \cdots A^{(i_r)}_{(m_r)}. \tag{11a}$$

Proof. We shall merely prove (11), the proof for (11a) being analogous. Let

$$T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = C^{(i_1)\cdots(i_r)}_{(m_1)\cdots(m_r)} A^{(m_1)}_{(j_1)} \cdots A^{(m_r)}_{(j_r)}. \tag{12}$$

It is readily verified that $T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$ satisfies the hypotheses of Lemma I. Hence $T^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = \lambda \cdot C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$. In particular, $\lambda \cdot C^{(i_1)\cdots(i_r)}_{(i_1)\cdots(i_r)} = T^{(i_1)\cdots(i_r)}_{(i_1)\cdots(i_r)}$. By (12), (6), and (10) this last equation may be written $\lambda \cdot r! = r!\, \Delta_C(A)$. Hence $\lambda = \Delta_C(A)$, and

$$\Delta_C(A) C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = C^{(i_1)\cdots(i_r)}_{(m_1)\cdots(m_r)} A^{(m_1)}_{(j_1)} \cdots A^{(m_r)}_{(j_r)}. \tag{12a}$$

On specializing in (12a) the superscripts of the rank tensor in accordance with (5a), we obtain (11).

From Lemma II, we obtain, corresponding to Theorem II of Part I, the

Cᴏʀᴏʟʟᴀʀʏ. $C_{(i_1)\cdots(i_r)}$ *and* $C^{(i_1)\cdots(i_r)}$ *are relative tensors of weights* w *and* $-w$ *respectively, where* $w = pr/n$.

The product theorem for bisymmetric tensors may be deduced from Lemma II:

$$\Delta_C(A)\cdot\Delta_C(B)=\Delta_C(P);\qquad P^{(i)}_{(j)}=A^{(i)}_{(m)}\,B^{(m)}_{(j)}. \tag{13}$$

The *adjoint* of $A^{(i)}_{(j)}$ *with respect to* $\mathfrak{S}$ is defined to be

$$\tilde{\mathfrak{A}}^{(i)}_{(j)}=\frac{1}{(r-1)\,!}\,C^{(i)\cdots(i_r)}_{(j)\cdots(j_r)}\,A^{(j_2)}_{(i_2)}\cdots A^{(j_r)}_{(i_r)}; \tag{14}$$

and the corresponding *inverse* is

$$\mathfrak{A}^{(i)}_{(j)}=\frac{1}{\Delta(A)}\cdot\tilde{\mathfrak{A}}^{(i)}_{(j)},\qquad (\Delta(A)\neq 0). \tag{15}$$

The proof of the fundamental properties

$$A^{(i)}_{(m)}\,\mathfrak{A}^{(m)}_{(j)}=\mathfrak{A}^{(i)}_{(m)}\,A^{(m)}_{(j)}=C^{(i)}_{(j)} \tag{16}$$

requires the identity

$$C^{(i_1)(i_2)\cdots(i_r)}_{(j_1)(i_2)\cdots(i_r)}=(r-1)\,!\,C^{(i_1)}_{(j_1)}. \tag{17}$$

Formula (17) is a special case of

$$C^{(i_1)\cdots(i_k)(i_{k+1})\cdots(i_r)}_{(j_1)\cdots(j_k)(i_{k+1})\cdots(i_r)}=(r-k)\,!\,C^{(i_1)\cdots(i_k)}_{(j_1)\cdots(j_k)}, \tag{18}$$

which we shall now prove. For any s we define

$$C^{(i_1)\cdots(i_s)}_{(j_1)\cdots(j_s)}=\begin{vmatrix} C^{(i_1)}_{(j_1)} & \cdots & C^{(i_1)}_{(j_s)}\\ \cdot & \cdots & \cdot\\ C^{(i_s)}_{(j_1)} & \cdots & C^{(i_s)}_{(j_s)}\end{vmatrix}. \tag{19}$$

Expanding this determinant by the last row, and rearranging indices, we get

$$\begin{aligned}C^{(i_1)\cdots(i_s)}_{(j_1)\cdots(j_s)}=&-C^{(i_s)}_{(j_1)}\,C^{(i_1)\cdots(i_{s-1})}_{(j_s)(j_2)\cdots(j_{s-1})}\\ &-C^{(i_s)}_{(j_2)}\,C^{(i_1)\cdots(i_{s-1})}_{(j_1)(j_s)(j_3)\cdots(j_{s-1})}-\cdots+C^{(i_s)}_{(j_s)}\,C^{(i_1)\cdots(i_{s-1})}_{(j_1)\cdots(j_{s-1})}.\end{aligned}$$

Thus, putting $(j_s)=(i_s)$, we have

$$C^{(i_1)\cdots(i_{s-1})(i_s)}_{(j_1)\cdots(j_{s-1})(i_s)}=\rho_s\cdot C^{(i_1)\cdots(i_{s-1})}_{(j_1)\cdots(j_{s-1})}, \tag{20}$$

where $\rho_s=C^{(i)}_{(i)}-(s-1)$. For $s=r+1$ the tensor (19) vanishes, and hence from (20),

$$0=\rho_{r+1}\cdot C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}.$$

The tensor on the right is not zero; therefore $\rho_{r+1}=C^{(i)}_{(i)}-r=0$, or

$$C^{(i)}_{(i)}=r. \tag{21}$$

Thus $\rho_s=r-s+1$, and by using (20) successively for $s=r, r-1,\cdots, k+1$ we obtain (18).

In equation (21) we have a fact of considerable importance; namely, that *the rank of an idempotent tensor* $C^{(i)}_{(j)}$ *is equal to* $C^{(i)}_{(i)}$. This corresponds to the well-known fact in matric algebra that the rank of an idempotent matrix is equal to its trace.

In this section we have restricted ourselves to a numerical idempotent tensor $C^{(i)}_{(j)}$ and to the algebra of bisymmetric tensors. If $A^{(i)}_{(j)}$ is not a bisymmetric tensor, we may still define the invariant (6); it is in fact the determinant of $C^{(i)}_{(l)} A^{(l)}_{(m)} C^{(m)}_{(j)}$,

$$\Delta_C(A) = \Delta_C(CAC).$$

We find however that

$$\Delta_C(A) \cdot \Delta_C(B) = \Delta_C(ACB);$$

this is a reflection of the fact that the mapping

$$A^{(i)}_{(j)} \rightarrow C^{(i)}_{(l)} A^{(l)}_{(m)} C^{(m)}_{(j)}$$

of the total $2p$-tensor algebra upon the algebra $\mathfrak{S}$ is not a homomorphism. Again, even if the idempotent $C^{(i)}_{(j)}$ is not numerical, the theory of this section will go through completely, with one exception, provided we restrict ourselves to an algebra of tensors commutative with $C^{(i)}_{(j)}$. The exception is that the entities $C^{(i_1)\cdots(i_r)}$ and $C_{(i_1)\cdots(i_r)}$ will not be tensors.

3. The direct sum of two algebras. Let the numerical tensors $D^{(i)}_{(j)}$ and $E^{(i)}_{(j)}$ be mutually orthogonal idempotents:

$$\begin{aligned} &D^{(i)}_{(m)} D^{(m)}_{(j)} = D^{(i)}_{(j)}, \qquad E^{(i)}_{(m)} E^{(m)}_{(j)} = E^{(i)}_{(j)}, \\ &D^{(i)}_{(m)} E^{(m)}_{(j)} = E^{(i)}_{(m)} D^{(m)}_{(j)} = 0. \end{aligned} \tag{22}$$

It follows from (22) that $C^{(i)}_{(j)} = D^{(i)}_{(j)} + E^{(i)}_{(j)}$ is a numerical idempotent tensor. Let r, s, t be the ranks of C, D, and E respectively. Then by contraction and by use of (21), $r = s + t$. Consider the rank tensor

$$C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = \begin{vmatrix} D^{(i_1)}_{(j_1)} + E^{(i_1)}_{(j_1)}, & \cdots, & D^{(i_1)}_{(j_r)} + E^{(i_1)}_{(j_r)} \\ \cdot & , \cdots, & \cdot \\ D^{(i_r)}_{(j_1)} + E^{(i_r)}_{(j_1)}, & \cdots, & D^{(i_r)}_{(j_r)} + E^{(i_r)}_{(j_r)} \end{vmatrix}. \tag{23}$$

The determinant on the right may be expressed as the sum of 2^r determinants of r rows and columns; the only non-zero ones will consist of s columns of D's and t columns of E's. Each such determinant may be expanded by minors, giving sums of products of s-rowed determinants of D's with t-rowed determinants of E's. Thus, finally, $C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$ may be expressed as an algebraic sum of $(r!/s!t!)^2$ products, of which the leading term may be taken to be

$$D^{(i_1)\cdots(i_s)}_{(j_1)\cdots(j_s)}\,E^{(i_{s+1})\cdots(i_r)}_{(j_{s+1})\cdots(j_r)}\,. \tag{24}$$

Let $\pi = \pi_1\pi_2\cdots\pi_s\pi_{s+1}\cdots\pi_r$ be any permutation of $1, 2, \cdots, r$ with the restriction that the numbers $\pi_1\pi_2\cdots\pi_s$ and also the numbers $\pi_{s+1}\cdots\pi_r$ are written in the natural order; let τ be another permutation of $1, 2, \cdots, r$ subject to the same restriction. Then a typical term of the sum whose leading term is (24) is

$$\pm\, D^{(i\pi_1)\cdots(i\pi_s)}_{(j\tau_1)\cdots(j\tau_s)}\,E^{(i\pi_{s+1})\cdots(i\pi_r)}_{(j\tau_{s+1})\cdots(j\tau_r)}, \tag{25}$$

where the sign is + or — according as π and τ differ by an even or an odd permutation. We note that there are $r!/s!t!$ terms for which $\pi = \tau$, all preceded by a plus sign.

THEOREM III. *For a bisymmetric tensor* $A^{(i)}_{(j)}$,

$$\Delta_C(A) = \Delta_D(A)\cdot\Delta_E(A). \tag{26}$$

Proof. In forming $\Delta_C(A)$ as defined by (6), where $C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$ is written as a sum as described above, if we consider only those terms for which $\tau = \pi$ we readily obtain the right hand side of (26). Any term for which $\tau \neq \pi$ will give zero. In fact, if $\tau \neq \pi$, one of the numbers $\pi_1\pi_2\cdots\pi_s$ must coincide with one of the numbers $\tau_{s+1}\tau_{s+2}\cdots\tau_r$; consider the case $\pi_1 = \tau_r = k$. The invariant corresponding to this term will vanish with

$$\begin{aligned} &D^{(ik)(i\pi_2)\cdots(i\pi_s)}_{(j\pi_1)(j\pi_2)\cdots(j\pi_s)}\cdot E^{(i\pi_{s+1})\cdots(i\pi_{r-1})(i\pi_r)}_{(j\pi_{s+1})\cdots(j\pi_{r-1})(jk)}\cdot A^{(jk)}_{(ik)} \\ &\qquad = D^{(ik)}_{(l)}\,D^{(l)(i\pi_2)\cdots(i\pi_s)}_{(j\tau_1)(j\tau_2)\cdots(j\tau_s)}\cdot E^{(i\pi_{s+1})\cdots(i\pi_{r-1})(i\pi_r)}_{(j\tau_{s+1})\cdots(j\tau_{r-1})(m)}\,E^{(m)}_{(jk)}\cdot A^{(jk)}_{(ik)}\,; \end{aligned}$$

for, since $A^{(i)}_{(j)}$ is bisymmetric,

$$D^{(ik)}_{(l)}\cdot E^{(m)}_{(jk)}\,A^{(jk)}_{(ik)} = D^{(ik)}_{(l)}\cdot A^{(m)}_{(jk)}\,E^{(jk)}_{(ik)} = A^{(m)}_{(jk)}\cdot E^{(jk)}_{(ik)}\,D^{(ik)}_{(l)} = 0,$$

the last step following from (22). A similar argument holds for the other cases.

Just as we have, in Theorem III, linked the determinants of $A^{(i)}_{(j)}$ with respect to C, D, and E, so we may link the adjoints ${}_C\bar{\mathfrak{A}}^{(i)}_{(j)}$, ${}_D\bar{\mathfrak{A}}^{(i)}_{(j)}$, ${}_E\bar{\mathfrak{A}}^{(i)}_{(j)}$, where these are defined relative to the idempotents C, D, and E in the manner indicated by formula (14).

THEOREM IV.

$${}_C\bar{\mathfrak{A}}^{(i)}_{(j)} = {}_D\bar{\mathfrak{A}}^{(i)}_{(j)}\cdot\Delta_E(A) + {}_E\bar{\mathfrak{A}}^{(i)}_{(j)}\cdot\Delta_D(A) \tag{27}$$

The proof of Theorem IV follows the same lines as the proof of Theorem III.

COROLLARY.

$$(28)\qquad {}_C\mathfrak{A}^{(i)}_{(j)} = {}_D\mathfrak{A}^{(i)}_{(j)} + {}_B\mathfrak{A}^{(i)}_{(j)}.$$

This follows from (27) by dividing through by $\Delta_C(A)$ and using (26).

Theorems III and IV may be immediately extended to the case when $C^{(i)}_{(j)}$ is a sum of any finite number of mutually orthogonal numerical idempotents. Such a situation arises when the unit tensor $\delta^{(i)}_{(j)}$ is decomposed into a (direct) sum of *immanant tensors,* as described in a previous paper by one of the writers [5]. Here

$$(29)\quad \delta^{i_1}_{j_1}\cdot\delta^{i_2}_{j_2}\cdots\delta^{i_p}_{j_p} = {}_{[p]}I^{i_1i_2\cdots i_p}_{j_1j_2\cdots j_p} + \cdots + {}_{[\alpha]}I^{i_1i_2\cdots i_p}_{j_1j_2\cdots j_p} + \cdots + {}_{[1^p]}I^{i_1i_2\cdots i_p}_{j_1j_2\cdots j_p},$$

or, more briefly,

$$(29\text{a})\qquad \delta^{(i)}_{(j)} = {}_{[p]}I^{(i)}_{(j)} + \cdots + {}_{[\alpha]}I^{(i)}_{(j)} + \cdots + {}_{[1^p]}I^{(i)}_{(j)}.$$

These formulas correspond to formula (18) in the paper cited. The immanant ${}_{[\alpha]}I^{(i)}_{(j)}$ is the idempotent numerical tensor corresponding to the irreducible representation $[\alpha]$ of the symmetric group on p letters. Like the idempotent $C^{(i)}_{(j)}$ of Section 2, the immanant ${}_{[\alpha]}I^{(i)}_{(j)}$ defines a subalgebra of the total $2p$-tensor algebra, and forms the unit for this subalgebra. Corresponding, respectively, to Theorem III and to the Corollary to Theorem IV we have, *for any bisymmetric tensor* $A^{(i)}_{(j)}$,

$$(30)\qquad |A| = \prod_{[\alpha]} \Delta_{[\alpha]}(A),$$

$$(31)\qquad \mathfrak{A}^{(i)}_{(j)} = \sum_{[\alpha]} {}_{[\alpha]}\mathfrak{A}^{(i)}_{(j)}.$$

As in Part I, $|A|$ and $\mathfrak{A}^{(i)}_{(j)}$ designate, respectively, the determinant and inverse of $A^{(i)}_{(j)}$ relative to $\delta^{(i)}_{(j)}$, that is, with respect to the total $2p$-tensor algebra. Clearly $\Delta_{[\alpha]}(A)$ and ${}_{[\alpha]}\mathfrak{A}^{(i)}_{(j)}$ designate the determinant and inverse, respectively, of $A^{(i)}_{(j)}$ relative to ${}_{[\alpha]}I^{(i)}_{(j)}$.

4. Determination of the ranks of immanants. If we contract in formula (29a) by putting $(j) = (i)$, we obtain

$$(32)\qquad \delta^{(i)}_{(i)} = {}_{[p]}I^{(i)}_{(i)} + \cdots + {}_{[\alpha]}I^{(i)}_{(i)} + \cdots + {}_{[1^p]}I^{(i)}_{(i)}.$$

In view of the fact given by formula (21), we may state (32) in words: *The rank of the unit tensor* ($= n^p$) *is equal to the sum of the ranks of the immanants.*

The immanant tensors were defined in terms of the table of characters of the symmetric group on p letters. We proceed to show how the ranks of these immanants may be obtained from the same table. If we define

$$(33) \qquad {}_{(\rho)}K^{(i)}_{(j)} = \sum_{\pi} \delta^{(i)}_{(j\pi)},$$

where π ranges over all permutations of class $(\rho) = (1^{\rho_1}, 2^{\rho_2}, \cdots, p^{\rho_p})$, then the immanant tensor ${}_{[\alpha]}I^{(i)}_{(j)}$ is defined by

$$(34) \qquad {}_{[\alpha]}I^{(i)}_{(j)} = (f\alpha/p!) \sum_{(\rho)} \chi_\alpha{}^{(\rho)} \cdot {}_{(\rho)}K^{(i)}_{(j)},$$

in which $\chi_\alpha{}^{(i)}$ is the characteristic for class (ρ) corresponding to the irreducible representation $[\alpha]$, and f_α is the degree of the representation $[\alpha]$, i. e., the characteristic corresponding to the class (1^p). If r_α is the rank of ${}_{[\alpha]}I^{(i)}_{(j)}$, we have from (34) by contraction,

$$(35) \qquad r_\alpha = (f\alpha/p!) \sum_{(\rho)} \chi_\alpha{}^{(\rho)} K^{(i)}_{(i)}.$$

LEMMA III.

$$(36) \qquad {}_{(\rho)}K^{(i)}_{(i)} = \nu_\rho \cdot n^k,$$

where ν_ρ = order of class (ρ), and $k \equiv k_\rho = \rho_1 + \rho_2 + \cdots \rho_p$.

Proof. If π is any permutation of class (ρ), $\delta^{(i)}_{(i\pi)} = \delta^{i_1}_{i\pi_1} \delta^{i_2}_{i\pi_2} \cdots \delta^{i_p}_{i\pi_p}$ is a product of closed cycles, ρ_1 of these being of length 1, ρ_2 of length 2, and so forth. Corresponding to a cycle of length λ, a factor of our product is

$$(37) \qquad \delta^{j_1}_{j_2} \delta^{j_2}_{j_3} \cdots \delta^{j_{\lambda-1}}_{j_\lambda} \delta^{j_\lambda}_{j_1} = \delta^{j_1}_{j_1} = n.$$

Hence $\delta^{(i)}_{(i\pi)} = n^k$, where $k = \rho_1 + \rho_2 + \cdots + \rho_p$. Since there are ν_ρ such permutations π, (36) follows.

From (35) and (36) there follows

THEOREM V.

$$(38) \qquad r_\alpha = (f_\alpha/p!) \sum_{(\rho)} \chi_\alpha{}^{(\rho)} \cdot \nu_\rho n^{k_\rho}.$$

Although the integers k_ρ are not usually given in character tables they are readily constructed from given character tables.

Example 1. For $p = 3$, the character table is [6]

Class: (ρ)	(1^3)	$(1, 2)$	(3)
Order: ν_ρ	1	3	2
k_ρ	3	2	1
[3]	1	1	1
[2, 1]	2	0	—1
$[1^3]$	1	—1	1

By (38), from the character table,

$$(39)\quad r_{[2,1]} = \frac{2}{3!}\,[2\cdot 1\cdot n^3 + 0\cdot 3\cdot n^2 - 1\cdot 2\cdot n] = \frac{4n(n^2-1)}{3!} = 4\cdot\binom{n+1}{3}.$$

In like manner it is found that $r_{[3]} = \binom{n+2}{3}$ and $r_{[1^3]} = \binom{n}{3}$.

• *Example* 2. For $p = 5$ we give only three lines of the character table, corresponding to class, order, and k_ρ.

Class	(1^5)	$(1^3, 2)$	$(1^2, 3)$	$(1, 4)$	$(1, 2^2)$	$(2, 3)$	(5)
Order	1	10	20	30	15	20	24
k_ρ	5	4	3	2	3	2	1

The ranks of the immanants are found to be

$[\alpha]$	$[5]$	$[4,1]$	$[3,2]$	$[3,1^2]$	$[2^2,1]$	$[2,1^3]$	$[1^5]$
$r_{[\alpha]}$	$\binom{n+4}{5}$	$16\cdot\binom{n+3}{5}$	$5n\cdot\binom{n+2}{4}$	$36\cdot\binom{n+2}{5}$	$5n\cdot\binom{n+1}{4}$	$16\cdot\binom{n+1}{5}$	$\binom{n}{5}$.

Since the immanants and their ranks are known, their rank tensors are readily constructed. As a simple instance we remark that for $n = p = 3$, $r_{[1^3]} = 1$, and so ${}_{[1^3]}I^{(i)}_{(j)} = \frac{1}{3!}\,\delta^{i_1 i_2 i_3}_{j_1 j_2 j_3}$ is its own rank tensor.

The $K^{(i)}_{(j)}$ defined by (33) is bisymmetric; in fact if $\kappa = \Sigma\pi$, the sum being over all permutations of class (ρ), and if τ, τ' are any permutation and its inverse, it readily follows that $\tau'\kappa\tau = \kappa$. Thus $\delta^{(i)}_{(l_{\tau'})}\,K^{(l)}_{(m)}\,\delta^{(m)}_{(j_\tau)} = K^{(i)}_{(j)}$, or in other words $K^{(i_\tau)}_{(j_\tau)} = K^{(i)}_{(j)}$, for every permutation τ; which proves the point. It then follows from (34) that the *immanant tensors are all bisymmetric.* Consider the set $\mathfrak{S}_{[\alpha]}$ of all tensors ${}_{[\alpha]}I^{(i)}_{(m)}\,A^{(m)}_{(j)} = A^{(i)}_{(m)}\,{}_{[\alpha]}I^{(m)}_{(j)}$ where $A^{(i)}_{(j)} \subset \mathfrak{B}$; clearly this is a linear subset of the algebra $\mathfrak{B}$. If $A^{(i)}_{(j)}$, $B^{(i)}_{(j)} \subset \mathfrak{B}$, then

$$B^{(i)}_{(l)}\cdot{}_{[\alpha]}I^{(l)}_{(m)}\,A^{(m)}_{(j)} = B^{(i)}_{(l)}\,A^{(l)}_{(m)}\cdot{}_{[\alpha]}I^{(m)}_{(j)},$$

and

$${}_{[\alpha]}I^{(i)}_{(m)}\,A^{(m)}_{(l)}\cdot B^{(l)}_{(j)} = {}_{[\alpha]}I^{(i)}_{(m)}\cdot A^{(m)}_{(l)}\,B^{(l)}_{(j)};$$

it follows that $\mathfrak{S}_{[\alpha]}$ *is an invariant subalgebra of* $\mathfrak{B}$.

5. The equivalent subalgebras connected with a Young tableau form. To facilitate our tensorial treatment of this topic, we need to extend the discussion of A. Young's work given in Section 1 of the paper "Tensor Algebra and Young's Symmetry Operators" [5].[2]

---

[2] The writers are indebted to Prof. G. de B. Robinson, in particular through his unpublished paper "The Irreducible Representations of the Symmetric Group, as obtained by Alfred Young," for helpful suggestions relative to this resumé.

Corresponding to a certain tableau form $[\alpha]$, let there be f_α standard forms, which we may designate by $F_1, F_2, \cdots, F_{f_\alpha}$. For each standard form there is a normalized idempotent symmetrizer $c_\rho = (f_\alpha/p\,!)P_\rho N_\rho$. For orthogonality purposes permutation operators M are introduced so that the *prepared* operators $d_\rho = c_\rho M_\rho$ have the properties $d_\rho d_\tau = 0$ or d_ρ according as $\rho \neq \tau$, $\rho = \tau$. Evidently $c_\rho d_\rho = c_\rho c_\rho M_\rho = c_\rho M_\rho = d_\rho$. Also, from the construction of the M_ρ, $N_\rho M_\rho \cdot P_\rho = N_\rho P_\rho$, whence there easily follows

$$P_\rho N_\rho M_\rho \cdot P_\rho N_\rho = P_\rho \cdot N_\rho M_\rho P_\rho \cdot N_\rho = P_\rho \cdot N_\rho P_\rho \cdot N_\rho = P_\rho N_\rho \cdot P_\rho N_\rho = \frac{p\,!}{f_\alpha} P_\rho N_\rho,$$

or $d_\rho c_\rho = c_\rho$. Thus

$$(40) \qquad \begin{aligned} c_\rho c_\rho &= c_\rho\,; & d_\rho d_\tau &= \delta_{\rho,\tau} \cdot d_\rho \\ c_\rho d_\rho &= d_\rho\,; & d_\rho c_\rho &= c_\rho, \end{aligned}$$

where $\delta_{\rho,\tau} = 1, 0$ according as $\rho = \tau$, $\rho \neq \tau$. Finally, if T is the normalized Young idempotent corresponding to the form of tableau $[\alpha]$, we know that

$$(41) \qquad T = d_1 + d_2 + \cdots + d_{f_\alpha}.$$

Dropping subscripts for the moment, let us assume that c, d are, respectively, the idempotent and the prepared idempotent corresponding to a certain standard form. Let $C^{(i)}_{(j)} \equiv C^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_p}$ be the numerical tensor corresponding to c under the homomorphic mapping (2), and similarly let $D^{(i)}_{(j)}$ be the numerical tensor corresponding to d. We deduce from (40) the relations

$$(42) \qquad \begin{aligned} C^{(i)}_{(m)}\, C^{(m)}_{(j)} &= C^{(i)}_{(j)}\,; & D^{(i)}_{(m)}\, D^{(m)}_{(j)} &= D^{(i)}_{(j)} \\ C^{(i)}_{(m)}\, D^{(m)}_{(j)} &= D^{(i)}_{(j)}\,; & D^{(i)}_{(m)}\, C^{(m)}_{(j)} &= C^{(i)}_{(j)}. \end{aligned}$$

THEOREM VI. *The idempotents $C^{(i)}_{(j)}$ and $D^{(i)}_{(j)}$ have the same rank, and $\Delta_D(A) = \Delta_C(A)$ for $A^{(i)}_{(j)}$ bisymmetric.*

Proof. Putting $(j) = (i)$ in the last two formulas of (42), we get

$$C^{(i)}_{(m)}\, D^{(m)}_{(i)} = D^{(i)}_{(i)}, \qquad D^{(i)}_{(m)}\, C^{(m)}_{(i)} = C^{(i)}_{(i)}.$$

Thus $D^{(i)}_{(i)} = C^{(i)}_{(i)}$, and the ranks are equal. We denote the common rank by r and the rank tensors by $D^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$ and $C^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)}$. From (42), $D^{(i)}_{(j)} = C^{(i)}_{(m)}\, D^{(m)}_{(j)}$, and

$$D^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = \begin{vmatrix} C^{(i_1)}_{(m_1)}\, D^{(m_1)}_{(j_1)}, & C^{(i_1)}_{(m_2)}\, D^{(m_2)}_{(j_2)}, & \cdots, & C^{(i_1)}_{(m_r)}\, D^{(m_r)}_{(j_r)} \\ \cdot & \cdot & \cdots & \cdot \\ C^{(i_r)}_{(m_1)}\, D^{(m_1)}_{(j_1)}, & C^{(i_r)}_{(m_2)}\, D^{(m_2)}_{(j_2)}, & \cdots, & C^{(i_r)}_{(m_r)}\, D^{(m_r)}_{(j_r)} \end{vmatrix}.$$

Since each (m_λ) ranges over $N = n^p$ labels, each column of the determinant on the right may be regarded as a sum of N columns. By expressing the determinant as the sum of the corresponding N^r determinants, we obtain

$$D^{(i_1)\cdots(i_r)}_{(j_1)\cdots(j_r)} = C^{(i_1)\cdots(i_r)}_{(m_1)\cdots(m_r)} \cdot D^{(m_1)}_{(j_1)} \cdots D^{(m_r)}_{(j_r)} . \tag{43}$$

For any bisymmetric tensor $A^{(i)}_{(j)}$ (43) yields, in condensed symbolism, $\Delta_D(A) = \Delta_C(DA)$. But $\Delta_C(DA) = \Delta_C(DA \cdot C) = \Delta_C(D \cdot CA) = \Delta_C(CA)$. Finally, $\Delta_D(A) = \Delta_C(A)$. Here D, C, A stand for $D^{(i)}_{(j)}, C^{(i)}_{(j)}, A^{(i)}_{(j)}$ and a product such as DA stands for $D^{(i)}_{(m)} A^{(m)}_{(j)}$.

Adopting the convention of H. Weyl [7], we shall use the name "natural diagram" in referring to the tableau, corresponding to a tableau form $[\alpha]$, in which the p numbers are written in the natural order. For example, if $p = 7$, the natural diagram corresponding to $[3, 2^2]$ is

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & & \end{array}$$

We shall designate by ϖ the diagram, of the same tableau form, in which the p numbers are written in an order obtained from the natural diagram with the aid of the permutation $\pi = \binom{1\,2\cdots p}{\pi_1\pi_2\cdots\pi_p}$. Let $C^{(i)}_{(j)} = C^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_p}$ be the numerical idempotent associated with the natural diagram, and ${}_{\varpi}C^{(i)}_{(j)}$ the numerical idempotent corresponding to the diagram ϖ. Denoting by π' the inverse permutation to π, we have

$${}_{\varpi}C^{(i)}_{(j)} = \delta^{(i)}_{(l\pi')} \cdot C^{(l)}_{(m)} \cdot \delta^{(m)}_{(j\pi)} = \delta^{(i\pi)}_{(l)} \cdot C^{(l)}_{(m)}\, \delta^{(m)}_{(j\pi)} = C^{i_{\pi_1} i_{\pi_2} \cdots i_{\pi_p}}_{j_{\pi_1} j_{\pi_2} \cdots j_{\pi_p}} = C^{(i_\pi)}_{(j_\pi)}. \tag{44}$$

Hence ${}_{\varpi}C^{(i)}_{(j)}$ *is obtained from* $C^{(i)}_{(j)}$ *by rearranging both subscripts and superscripts of the latter according to the permutation* π.

If $A^{(j)}_{(j)}$ is any bisymmetric tensor, consider the set $\mathfrak{U}$ of all tensors of the form $U^{(i)}_{(j)} = C^{(i)}_{(m)} A^{(m)}_{(j)}$. Although $\mathfrak{U}$ is a subalgebra of the total $2p$-tensor algebra, it is not a subalgebra of the algebra $\mathfrak{B}$ of bisymmetric tensors. In fact

$$U^{(i\pi)}_{(j\pi)} = C^{(i\pi)}_{(m)} A^{(m)}_{(j\pi)} = C^{(i\pi)}_{(m\pi)} A^{(m\pi)}_{(j\pi)} = {}_{\varpi}C^{(i)}_{(m)} A^{(m)}_{(j)} = {}_{\varpi}U^{(i)}_{(j)},$$

the last tensor belonging to the analogously formed algebra ${}_{\varpi}\mathfrak{U}$.

The mapping

$$U^{(i)}_{(j)} \leftrightarrow {}_{\varpi}U^{(i)}_{(j)} = U^{(i\pi)}_{(j\pi)} \tag{45}$$

is an isomorphic mapping of $\mathfrak{U}$ *upon* ${}_{\varpi}\mathfrak{U}$, *and hence these algebras are equivalent.* Here we have a situation in contrast with that which arises in the case of the corresponding algebra defined by an immanant; the latter is a subalgebra, in fact an invariant subalgebra, of the algebra $\mathfrak{B}$.

From the remarks of the preceding paragraph it is evident that once the inverse has been formed relative to C, it may be formed relative to ${}_{\varpi}C$ by use of the permutation π. This is a consequence of the fact that the mapping (45) is an isomorphism.

THEOREM VII. *The determinant of a bisymmetric tensor* $A^{(i)}_{(j)}$ *relative to* ${}_{\varpi}C^{(i)}_{(j)}$ *is equal to its determinant relative to* $C^{(i)}_{(j)}$; *that is,*

$$\Delta_{\varpi C}(A) = \Delta_C(A). \tag{46}$$

The proof is left to the reader.

Corresponding to equation (41) we may write

$${}_{[\alpha]}I^{(i)}_{(j)} = D_1{}^{(i)}_{(j)} + D_2{}^{(i)}_{(j)} + \cdots + D_{f_\alpha}{}^{(i)}_{(j)}. \tag{47}$$

Here ${}_{[\alpha]}I$ is the immanant associated with the tableau form $[\alpha]$, and the D's are the prepared (mutually orthogonal) tensor idempotents. We have directly from Theorem III

$$\Delta_I(A) = \Delta_{D_1}(A) \cdot \Delta_{D_2}(A) \cdots \Delta_{D_{f_\alpha}}(A). \tag{48}$$

This, with Theorems VI and VII, allows us to state the

COROLLARY. $\Delta_I(A) = [\Delta_C(A)]^{f_\alpha}$, *where* $C^{(i)}_{(j)}$ *is the idempotent for the natural diagram of the tableau form for which* $I^{(i)}_{(j)}$ *is the immanant.*

Like the algebra $\mathfrak{U}$ of tensors $U^{(i)}_{(j)} = C^{(i)}_{(m)}A^{(m)}_{(j)}$ ($A^{(i)}_{(j)}$ bisymmetric) the algebra $\mathfrak{V}$ of tensors $V^{(i)}_{(j)} = D^{(i)}_{(m)}A^{(m)}_{(j)}$ forms a subalgebra of the total $2p$-tensor algebra, but not of the bisymmetric algebra $\mathfrak{B}$. It is interesting to note that, as implied in (47), the algebra $\mathfrak{S}_{[\alpha]}$ considered in the preceding section, which is an invariant subalgebra of $\mathfrak{B}$, is the direct sum of f_α algebras $\mathfrak{V}$, none of which is a subalgebra of $\mathfrak{B}$:

THEOREM VIII. *Each of the algebras* $\mathfrak{V}_1, \mathfrak{V}_2, \cdots, \mathfrak{V}_{f_\alpha}$ *is equivalent to the algebra* $\mathfrak{U}$ *defined with respect to the idempotent* $C^{(i)}_{(j)}$ *of the natural diagram.*

Proof. Let $D^{(i)}_{(j)}$ be the prepared idempotent corresponding to the natural diagram. The algebra $\mathfrak{U}$ consists of all tensors of the form $U^{(i)}_{(j)} = C^{(i)}_{(m)}A^{(m)}_{(j)}$, $A^{(i)}_{(j)}$ bisymmetric. Define

$$\bar{U}^{(i)}_{(j)} = D^{(i)}_{(k)}\, U^{(k)}_{(l)}\, D^{(l)}_{(j)}.$$

Then

$$\begin{aligned} \bar{U}^{(i)}_{(j)} = D^{(i)}_{(k)} \cdot C^{(k)}_{(m)} A^{(m)}_{(l)} \cdot D^{(l)}_{(j)} &= D^{(i)}_{(k)} C^{(k)}_{(m)} \cdot D^{(m)}_{(l)} \cdot A^{(l)}_{(j)} && (A \text{ bisymmetric}) \\ &= C^{(i)}_{(m)} \cdot D^{(m)}_{(l)} \cdot A^{(l)}_{(j)}. && (\text{By } (42)) \end{aligned}$$

So

$$\bar{U}^{(i)}_{(j)} = V^{(i)}_{(j)} \equiv D^{(i)}_{(l)}\, A^{(l)}_{(j)}. \qquad (\text{By } (42)) \tag{49}$$

Conversely

(50) $$\bar{V}^{(i)}_{(j)} \equiv C^{(i)}_{(k)}\, V^{(k)}_{(l)}\, C^{(l)}_{(j)} = U^{(i)}_{(j)}.$$

The mapping $U^{(i)}_{(j)} \rightarrow D^{(i)}_{(k)}\, U^{(k)}_{(l)}\, D^{(l)}_{(j)} = V^{(i)}_{(j)}$ is a homomorphic mapping of $\mathfrak{U}$ upon another algebra, which, as we see from (49), is none other than the algebra $\mathfrak{B}$ defined in the preceding paragraph. From (50) we deduce that the mapping is in fact an isomorphism: $\mathfrak{B}_1$ is equivalent to $\mathfrak{U}_1$. Similarly we may show $\mathfrak{B}_\lambda$ to be equivalent to $\mathfrak{U}_\lambda$ for $\lambda = 2, 3, \cdots, f_a$. Since all the algebras $\mathfrak{U}_\lambda$ $(\cong \mathfrak{U})$ are equivalent to $\mathfrak{U}$ the theorem follows.

6. Conclusion. In Section 2 we showed that in terms of any idempotent numerical tensor $C^{(i)}_{(j)}$ we might define a subalgebra of the total $2p$-tensor algebra and that $C^{(i)}_{(j)}$ was the unit of this algebra and had rank $C^{(i)}_{(i)}$. In terms of the rank tensor of $C^{(i)}_{(j)}$ we were able to define the determinant, and the inverse when this existed, of any member of the subalgebra. These general considerations facilitated the study of the invariant subalgebras $\mathfrak{S}_{[a]}$ of the algebra $\mathfrak{B}$ of bisymmetric tensors; $\mathfrak{S}_{[a]}$ being defined in terms of the immanant tensor ${}_{[a]}I^{(i)}_{(j)}$. In Section 4 there was given a method for the explicit calculation of the rank of any immanant; explicit construction of the determinant and inverse of any tensor belonging to $\mathfrak{S}_{[a]}$ was thus made possible. Section 5 established the fact that $\mathfrak{S}_{[a]}$ was the direct sum of f_a equivalent tensor algebras, none of which, however, was a subalgebra of $\mathfrak{B}$.

It seems pertinent that we conclude our theory with

THEOREM IX. *The algebras $\mathfrak{S}_{[a]}$ are the only minimal invariant subalgebras of $\mathfrak{B}$.*

Proof. Let $\mathfrak{S}$, with unit tensor $I^{(i)}_{(j)}$, be a minimal invariant subalgebra of $\mathfrak{B}$. The tensor $I^{(i)}_{(j)}$ is bisymmetric, since it belong to $\mathfrak{B}$, and idempotent, since it is the unit of $\mathfrak{S}$. For any $U^{(i)}_{(j)} \subset \mathfrak{S}$, we have $U^{(i)}_{(m)}\, I^{(m)}_{(j)} = I^{(i)}_{(m)}\, U^{(m)}_{(j)} = U^{(i)}_{(j)}$, from the properties of a unit. If $A^{(i)}_{(j)} \subset \mathfrak{B}$, then, since $\mathfrak{S}$ is an invariant subalgebra of $\mathfrak{B}$, $A^{(i)}_{(m)}\, I^{(m)}_{(j)} \subset \mathfrak{S}$, and hence $A^{(i)}_{(m)}\, I^{(m)}_{(j)} = A^{(i)}_{(m)}\, I^{(m)}_{(l)} \cdot I^{(l)}_{(j)} = I^{(i)}_{(l)} \cdot A^{(l)}_{(m)}\, I^{(m)}_{(j)} = I^{(i)}_{(l)}\, A^{(l)}_{(m)}\, I^{(m)}_{(j)}$. Similarly, $I^{(i)}_{(l)}\, A^{(l)}_{(j)} = I^{(i)}_{(l)}\, A^{(l)}_{(m)}\, I^{(m)}_{(j)}$; whence

(51) $$A^{(i)}_{(m)}\, I^{(m)}_{(j)} = I^{(i)}_{(l)}\, A^{(l)}_{(j)}.$$

Equation (51) states that the idempotent $I^{(i)}_{(j)}$ lies in the commutator algebra, $\mathfrak{L}$, of $\mathfrak{B}$; i. e., it is numerical (see Theorem II) and thus a linear sum of the tensors $\delta^{(i)}_{(j_\pi)}$.

Let us first consider the case $n \geqq p$, so that the mapping (2) of the group ring $\mathfrak{S}_p$ upon the algebra $\mathfrak{L}$ is an isomorphism, or, in other words, so that the

$p!$ tensors $\delta^{(i)}_{(j\pi)}$ are linearly independent and $I^{(i)}_{(j)}$ may be expressed as a linear sum of these tensors in a unique manner. But $I^{(i)}_{(j)}$ is bisymmetric; hence the coefficient of each $\delta^{(i)}_{(j)}$ for which π is of class (ρ) must be the same; in other words, $I^{(i)}_{(j)}$ is a sum of the "class-tensors" ${}_{(\rho)}K^{(i)}_{(j)}$ defined by equation (33). We now use the well-known fact that the number of irreducible representations of the symmetric group is equal to the number of classes. Since the linearly independent immanants ${}_{[a]}I^{(i)}_{(j)}$ are equal in number to the "class tensors" ${}_{(\rho)}K^{(i)}_{(j)}$ in terms of linear sums of which they are defined, it follows that the ${}_{(\rho)}K^{(i)}_{(j)}$ can be written as a linear sum of the immanants. Hence the idempotent $I^{(i)}_{(j)}$ is a linear sum of the (mutually orthogonal) immanant tensors, and each coefficient in the sum will necessarily be 1 or 0. But $\mathfrak{S}$ is minimal. Therefore $I^{(i)}_{(j)}$ is itself an immanant; $\mathfrak{S} = \mathfrak{S}_{[a]}$.

If $n < p$ the above remarks must be re-examined, but the theorem is nevertheless true. The class-tensors ${}_{(\rho)}K^{(i)}_{(j)}$ and the immanants ${}_{[a]}I^{(i)}_{(j)}$ are defined by (33) and (34) respectively as the "homomorphs" under (2) of definite group-ring elements; hence the truth of the proposition that the ${}_{(\rho)}K^{(i)}_{(j)}$ are expressible linearly in terms of the ${}_{[a]}I^{(i)}_{(j)}$ depends only upon the group-ring $\mathfrak{S}_p$ and not upon the relation of n to p. The linear subset $\mathfrak{L}$ consisting of those elements of $\mathfrak{S}_p$ which map under (2) into the zero tensor is clearly an invariant sub-algebra of $\mathfrak{S}_p$, and $\mathfrak{S}_p = \mathfrak{L} \oplus \mathfrak{M}$; the group-ring is a direct sum of $\mathfrak{L}$ and another invariant subalgebra $\mathfrak{M}$. Of the various group-ring operators which map into $I^{(i)}_{(j)}$ we select that (uniquely determined) operator ϵ which lies in $\mathfrak{M}$. (The tensor $I^{(i)}_{(j)}$ may be written in different forms as a linear sum of the isomers; with the aid of the group-ring we select a definite one of these expressions.) Thus $\epsilon \to I^{(i)}_{(j)}$. If π, π' are any permutation and its inverse, $\pi' \epsilon \pi \to \delta^{(i)}_{(l\pi')} \cdot I^{(l)}_{(m)} \cdot \delta^{(m)}_{(j\pi)} = I^{(i\pi)}_{(j\pi)}$. Thus, since $I^{(i)}_{(j)}$ is bisymmetric, $\pi' \epsilon \pi \to I^{(i)}_{(j)}$. Now $\mathfrak{M}$ is an invariant subalgebra, and so $\pi' \epsilon \pi$ lies in $\mathfrak{M}$, and both $\pi' \epsilon \pi$ and ϵ, members of $\mathfrak{M}$, map into $I^{(i)}_{(j)}$. Therefore $\pi' \epsilon \pi = \epsilon$ for all π; or; ϵ lies in the centrum of $\mathfrak{S}_p$. It follows that $I^{(i)}_{(j)}$ is linearly expressible in terms of the ${}_{(\rho)}K^{(i)}_{(j)}$ and hence in terms of the ${}_{[a]}I^{(i)}_{(j)}$. We conclude, as in the preceding paragraph, that $I^{(i)}_{(j)}$ is itself an immanant.

7. Applications.

Of aid in studying the algebra of tensors associated with a Young diagram is the

LEMMA IV. *Let* $C^{(i)}_{(j)}$ *be the numerical idempotent tensor associated with the natural diagram of a Young tableau form. Then* $C^{(i)}_{(j)} \equiv C^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_p}$ *possesses:*

(1) *on its p superscripts the symmetries indicated by the rows of the diagram;*

(2) *on its p subscripts the skew-symmetries indicated by the columns of the diagram.*

Remark. Here the idempotent $C^{(i)}_{(j)}$ is formed with the aid of an idempotent symmetrizer $c = (f_a/p!)PN$; if the order of P and N were interchanged (multiplication of permutations still being defined as in section 1), one would interchange the words "superscripts" and "subscripts" in the Lemma.

Illustration. The idempotent $C^{(i)}_{(j)} \equiv C^{i_1 i_2 i_3 i_4 i_5 i_6 i_7}_{j_1 j_2 j_3 j_4 j_5 j_6 j_7}$ associated with the diagram given in Section **5** is: (1) symmetric on i_1, i_2, i_3, on i_4, i_5, i_6 and on i_7; (2) skew-symmetric on j_1, j_4, j_7, on j_2, j_5 and on j_3, j_6.

Proof. $C^{(i)}_{(j)}$ is obtained by applying the group-ring operation $c = (f_a/p!)PN$ to the subscripts of the unit tensor $\delta^{(i)}_{(j)}$. Here P is the sum of the admissible row permutations; thus if π is any row permutation, $\pi P = P$ and so $\pi c = c$. The last equation, in virtue of the homomorphic correspondence between the group ring and the algebra of absolute numerical tensors, implies $\delta^{(i)}_{(m\pi)} C^{(m)}_{(j)} = C^{(i)}_{(j)}$. But $\delta^{(i)}_{(m\pi)} C^{(m)}_{(j)} = \delta^{(i\pi')}_{(m)} C^{(m)}_{(j)} = C^{(i\pi')}_{(j)}$, where π' is the inverse permutation to π; hence $C^{(i\pi')}_{(j)} = C^{(i)}_{(j)}$. Since π' may range with π over all admissible row permutations, we infer property (1) of the lemma. On the other hand, N is the alternating sum of the admissible column permutations; if ν is any column permutation, $N \cdot \nu = \lambda_\nu \cdot N$ where $\lambda_\nu = +1, -1$ according as the permutation ν is even or odd. Thus $c \cdot \nu = \lambda_\nu \cdot c$, which implies $C^{(i)}_{(m)} \cdot \delta^{(i)}_{(j\nu)} = C^{(i)}_{(j)}$, or $C^{(i)}_{(j\nu)} = C^{(i)}_{(j)}$. Hence we have property (2).

The fact that $C^{(i)}_{(j)}$ is associated with the natural diagram plays no part in our proof. Lemma IV would remain true if the words "the natural" were replaced by the indefinite article. Nothing essential is lost, however, by our keeping to the natural diagram; this we have shown in Section **5**.

Example: $p = 3$, $[\alpha] = [2, 1]$, $n = 2$. Here $f_a = 2$; there are two standard diagrams

$$F_1: \begin{matrix} 12; \\ 3 \end{matrix} \qquad F_2: \begin{matrix} 13 \\ 2 \end{matrix}$$

From Example 1, Section **4**, $r_{[2,1]} = 4 \cdot \binom{n+1}{3}$ for general n. In the present case $r_{[2,1]} = 4$, and the rank of each of the idempotents $C_1{}^{(i_1)}_{(j_1)}$, $C_2{}^{(i_1)}_{(j_1)}$ associated with F_1, F_2 is $(1/f_a) \cdot r_a = 2$.

From Lemma IV, $C_1{}^{(i)}_{(j)} \equiv C_1{}^{i_1 i_2 i_3}_{j_1 j_2 j_3}$ is symmetric in the set $i_1\, i_2$ and skew-symmetric in the set $j_1\, j_3$. The rank tensor of $C_1{}^{(i)}_{(j)}$ is

$$C_1{}^{(i_1)(i_2)}_{(j_1)(j_2)} = \begin{vmatrix} C_1{}^{i_{11} i_{12} i_{13}}_{j_{11} j_{12} j_{13}}, & C_1{}^{i_{11} i_{12} i_{13}}_{j_{21} j_{22} j_{23}} \\ C_1{}^{i_{21} i_{22} i_{23}}_{j_{11} j_{12} j_{13}}, & C_1{}^{i_{21} i_{22} i_{23}}_{j_{21} j_{22} j_{23}} \end{vmatrix}.$$

From its construction this tensor is symmetric on $i_{\lambda_1} i_{\lambda_2}$ and skew-symmetric on $j_{\lambda_1} j_{\lambda_3}$ for $\lambda = 1, 2$. In addition it is skew-symmetric in $(i_1)(i_2)$ and in $(j_1)(j_2)$. (See Lemma I). The rank tensor, moreover, has rank 1 (Corollary to Theorem I, Part I) and hence is unique (See Lemma I). Also it factors: $C_1{}^{(i_1)(i_2)}_{(j_1)(j_2)} = (1/c) \cdot C_1{}^{(i_1)(i_2)} \cdot C_{1\,(j_1)(j_2)}$ for some scalar c. Consider the tensor $\epsilon_{j_{11} j_{13}} \cdot \epsilon_{j_{21} j_{23}} \cdot \epsilon_{j_{12} j_{22}}$. This is skew-symmetric in $j_{\lambda_1} j_{\lambda_3}$ for $\lambda = 1, 2$; if we interchange the rôles of (j_1) and (j_2) we merely change the sign of the tensor. Hence

$$C_{1\,(j_1)(j_2)} = \epsilon_{j_{11} j_{13}} \cdot \epsilon_{j_{21} j_{23}} \cdot \epsilon_{j_{12} j_{22}} \tag{52}$$

apart from a constant factor. In order to obtain $C_1{}^{(i_1)(i_2)}$, we remark that the corresponding tensor $\epsilon^{i_{11} i_{13}} \cdot \epsilon^{i_{21} i_{23}} \cdot \epsilon^{i_{12} i_{22}}$ is skew-symmetric in $(i_1)(i_2)$ and also in $i_{\lambda_1} i_{\lambda_3}$ for $\lambda = 1, 2$. We may obtain $C_1{}^{(i_1)(i_2)}$ from this tensor by applying $N_1 P_1$ independently to the three indices (i_1) and to the three indices (i_2); it is unnecessary however to apply N_1, which will have the effect of multiplying the tensor by a scalar. Applying $P_1 = \{12\} \cdot \{3\} = 1 + (12)$ we get

$$\begin{aligned} C_1{}^{(i_1)(i_2)} = {} & \epsilon^{i_{11} i_{13}} \cdot \epsilon^{i_{21} i_{23}} \cdot \epsilon^{i_{12} i_{22}} \\ & + \epsilon^{i_{12} i_{13}} \cdot \epsilon^{i_{21} i_{23}} \cdot \epsilon^{i_{11} i_{22}} + \epsilon^{i_{11} i_{13}} \cdot \epsilon^{i_{22} i_{23}} \cdot \epsilon^{i_{12} i_{21}} + \epsilon^{i_{12} i_{13}} \cdot \epsilon^{i_{22} i_{23}} \cdot \epsilon^{i_{11} i_{21}} \end{aligned} \tag{53}$$

apart from a constant factor. Outer multiplication of the expressions (52) and (53) for $C_{1(j_1)(j_2)}$ and $C_1{}^{(i_1)(i_2)}$ yields

$$\begin{aligned} C_1{}^{(i_1)(i_2)}_{(j_1)(j_2)} = \lambda \cdot [& \delta^{i_{11} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{21} i_{23}}_{j_{21} j_{23}} \cdot \delta^{i_{12} i_{22}}_{j_{12} j_{22}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{21} i_{23}}_{j_{21} j_{23}} \cdot \delta^{i_{11} i_{22}}_{j_{12} j_{22}} \\ & + \delta^{i_{11} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{22} i_{23}}_{j_{21} j_{23}} \cdot \delta^{i_{12} i_{21}}_{j_{12} j_{22}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{22} i_{23}}_{j_{21} j_{23}} \cdot \delta^{i_{11} i_{21}}_{j_{12} j_{22}}] \end{aligned} \tag{54}$$

for some scalar λ to be determined. If we contract by putting $(j_2) = (i_2)$, formula (18) where $r = 2$ gives $C_1{}^{(i_1)(i_2)}_{(j_1)(j_2)} = 2 \cdot C_1{}^{(i_1)}_{(j_1)}$. Hence from (54),

$$\begin{aligned} C_1{}^{(i_1)}_{(j_1)} & = \frac{\lambda}{2} [\delta^{i_{11} i_{13}}_{j_{11} j_{13}} \cdot 2 \cdot \delta^{i_{12}}_{j_{12}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}} \cdot 2 \cdot \delta^{i_{11}}_{j_{12}} + \delta^{i_{11} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{12}}_{j_{12}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}} \cdot \delta^{i_{11}}_{j_{12}}] \\ & = \frac{3\lambda}{2} [\delta^{i_{12}}_{j_{12}} \cdot \delta^{i_{11} i_{13}}_{j_{11} j_{13}} + \delta^{i_{11}}_{j_{12}} \cdot \delta^{i_{12} i_{13}}_{j_{11} j_{13}}]. \end{aligned} \tag{55}$$

Finally, since $C_1{}^{(i_1)}_{(i_1)} = 2$, $2 = \frac{3\lambda}{2} \cdot [2 \cdot 2 + 2] = 9\lambda$, $\lambda = \frac{2}{9}$. Hence, dropping the subscript 1 on (i_1), (j_1), we may write

$$(56) \qquad C_1{}^{(i)}_{(j)} = \frac{1}{3}\left[\delta^{i_2}_{j_2}\cdot\delta^{i_1 i_3}_{j_1 j_3} + \delta^{i_1}_{j_2}\cdot\delta^{i_2 i_3}_{j_1 j_3}\right] = \frac{1}{3}\left[\delta^{i_2}_{j_2}\cdot\epsilon^{i_1 i_3} + \delta^{i_1}_{j_2}\cdot\epsilon^{i_2 i_3}\right]\cdot\epsilon_{j_1 j_3}.$$

Also

$$C_1{}^{(i_1)(i_2)}_{(j_1)(j_2)} = \frac{2}{9}\left[\delta^{i_{11} i_{13}}_{j_{11} j_{13}}\cdot\delta^{i_{21} i_{23}}_{j_{21} j_{23}}\cdot\delta^{i_{12} i_{22}}_{j_{12} j_{22}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}}\cdot\delta^{i_{21} i_{23}}_{j_{21} j_{23}}\cdot\delta^{i_{11} i_{22}}_{j_{12} j_{22}}\right.$$
$$\left. + \delta^{i_{11} i_{13}}_{j_{11} j_{13}}\cdot\delta^{i_{22} i_{23}}_{j_{21} j_{23}}\cdot\delta^{i_{12} i_{21}}_{j_{12} j_{22}} + \delta^{i_{12} i_{13}}_{j_{11} j_{13}}\cdot\delta^{i_{22} i_{23}}_{j_{21} j_{23}}\cdot\delta^{i_{11} i_{21}}_{j_{12} j_{22}}\right].$$

The idempotent $C_2{}^{(i)}_{(j)}$ is obtained from $C_1{}^{(i)}_{(j)}$ as given by (56) by use of the permutation $\pi = (23)$:

$$(57) \qquad C_2{}^{(i)}_{(j)} = \frac{1}{3}\left[\delta^{i_3}_{j_3}\cdot\delta^{i_1 i_2}_{j_1 j_2} - \delta^{i_1}_{j_3}\cdot\delta^{i_2 i_3}_{j_1 j_2}\right] = \frac{1}{3}\left[\delta^{i_3}_{j_3}\cdot\epsilon^{i_1 i_2} - \delta^{i_1}_{j_3}\cdot\epsilon^{i_2 i_3}\right]\cdot\epsilon_{j_1 j_2}.$$

It is readily verified that the idempotents $C_1{}^{(i)}_{(j)}$, $C_2{}^{(i)}_{(j)}$ are mutually orthogonal; and in fact both M_1 and M_2 are here the identical permutation. Thus ${}_{[2,1]}I^{(i)}_{(j)} = C_1{}^{(i)}_{(j)} + C_2{}^{(i)}_{(j)}$. From (56) and (57), using the identity $\delta^{i_1}_{j_1}\cdot\epsilon_{j_2 j_3} - \delta^{i_1}_{j_2}\cdot\epsilon_{j_1 j_3} + \delta^{i_1}_{j_3}\cdot\epsilon_{j_1 j_2} = 0$, we obtain

$$(58) \qquad {}_{[2,1]}I^{(i)}_{(j)} = \frac{1}{3}\left[\delta^{i_1}_{j_1}\cdot\delta^{i_2 i_3}_{j_2 j_3} + \delta^{i_2}_{j_2}\cdot\delta^{i_1 i_3}_{j_1 j_3} + \delta^{i_3}_{j_3}\cdot\delta^{i_1 i_2}_{j_1 j_2}\right].$$

For theoretical purposes the idempotent associated with a Young diagram is defined as the result of operating on the unit tensor $\delta^{(i)}_{(j)}$ with a Young symmetrizer. Nevertheless, the above example has served to indicate the truth of

LEMMA V. *The idempotent associated with a Young diagram is uniquely defined by its rank and the symmetries indicated by the diagram; and it may be expressed in terms of the (ordinary) generalized Kronecker deltas.* (See (23), Part I).

In fact, the rank tensor is determined to within a constant by its symmetries and hence may be determined to within a constant in terms of the generalized Kronecker delta $\delta^{i_1 i_2 \cdots i_n}_{j_1 j_2 \cdots j_n}$. Both the constant and the idempotent may be obtained from the rank tensor by use of the contraction process (18).

Although the immanant may be obtained if desired from the character table as a function of the simple Kronecker delta, the example indicates how it may be expressed in a more manageable form through the C's by the use of Lemma V.

In the Example the M's, as there noted, were each the identity permutation. This is true in all cases for $p < 5$. In some cases, as for example when $p = 5$, $[\alpha] = [3, 2]$ and $[\alpha] = [2^2, 1]$, some M's are not the identity and must be determined. For determination of the M's see A. Young [8].

UNIVERSITY OF ALABAMA.

REFERENCES

[1] A. Adrian Albert, *Modern Higher Algebra*, Chicago (1937), p. 139.

[2] H. Weyl, *The Classical Groups*, Princeton (1939), p. 119.

[3] C. M. Cramlet, "A determination of all invariant tensors," *Tohoku Mathematical Journal*, vol. 28 (1927), pp. 242-250.

[4] M. S. Knebelman, "Tensors with invariant components," *Annals of Mathematics*, 2, vol. 30 (1929), pp. 339-344.

[5] T. L. Wade, "Tensor algebra and Young's symmetry operators," *American Journal of Mathematics*, vol. 63 (1941), pp. 645-657.

[6] D. E. Littlewood and A. R. Richardson, "Group characters and algebra," *Philosophical Transactions of the Royal Society of London*, (A), vol. 233 (1934), pp. 99-141.

[7] H. Weyl, *The Theory of Groups and Quantum Mechanics*, New York, 1932, p. 360.

[8] A. Young, "On quantitative substitutional analysis," *The Journal of the London Mathematical Society*, vol. 3 (1928), pp. 14-19.

Note on a Paper by R. H. Bruck and T. L. Wade.

By G. de B. Robinson.

In Theorem V, Part II, of their paper on 'Bisymmetric Tensor Algebra' R. H. Bruck and T. L. Wade have expressed the *rank* r_a of an 'immanant tensor' ${}_{[a]}I^{(i)}_{(j)}$ by the following formula:

$$r_a = (f_a/p\,!) \sum_{(\rho)} \chi_a{}^{(\rho)} \cdot \nu_\rho n^{k_\rho}, \tag{1}$$

where f_a = degree of the representation $[\alpha]$ of S_p;
$\chi_a{}^{(\rho)}$ = characteristic of the ρ-th conjugate set in $[\alpha]$;
ν_p = number of elements of the ρ-th conjugate set;
k_p = number of cycles of an element of the ρ-th conjugate set.

It is perhaps worth pointing out an interesting interpretation for the number r_a in terms of Young's substitutional analysis. If we replace each cycle in the fundamental relation *

$$T_a = (f_a/p\,!) \sum_P \chi_a{}^{P^{-1}} \cdot P = (f_a/p\,!)^2 \sum P_a N_a \tag{2}$$

by n, then immediately

$$r_a = (f_a/p\,!) \sum_{(\rho)} \chi_a{}^{(\rho)} \cdot \nu_\rho n^{k_\rho} = (f_a/p\,!)^2 \sum \{P_a N_a\} = (f_a{}^2/p\,!)\{PN\},$$

since the second summation in (2) is over all the $p\,!$ tableaux α and the corresponding $\{PN\}$'s—i. e. PN with each cycle replaced by n—are all equal.

This interpretation adds to the significance of the table in Example 2 of Section **4**. E. g.

$$r_{[5]} = \frac{1}{5\,!}\,(n^5 + 10n^4 + (20+15)n^3 + (30+20)n^2 + 24n) = \binom{n+4}{5},$$

$$r_{[3,2]} = \frac{5^2}{5\,!}\,(n^5 + 2n^4 + (-4+3)n^3 + (-6+4)n^2) = 5n\binom{n+2}{4}, \text{ etc.}$$

as may be easily verified.

---

* Cf. Robinson, "On the Representations of the Symmetric Group" this Journal, vol. 60 (1938), pp. 745-760, footnote 12.

QUALITATIVE ANALYSIS OF THE FLOW OF FLUIDS IN PIPES.*

By T. Y. Thomas.

1. Introduction. In another paper: *On the stability of viscous fluids*[1] we obtained a number of results on stability problems belonging to the first category in the classification of this paper. Stability problems of the second category, e. g., the stability of steady flow (Poiseuille motion) in an infinite circular cylinder or pipe, were not considered. This particular problem will be discussed in the present paper, the methods and the chief results obtained having their analogues in other stability problems treated in the literature.

Exact methods only are used in this paper, e. g., the linearized or first order equations of the disturbance are not considered.[2] While it is natural to consider exponential solutions (disturbances) in connection with the linearized equations it is questionable if such disturbances (excepting the special case of the zero disturbance) have an actual existence in the sense that they satisfy the exact equations of disturbance. At the end we have in fact shown that the procedure of representing disturbances by functions of exponential type leads (on the basis of the exact equations) to the stability of the steady flow, a result in disagreement with experiment. Other results on stability are mentioned in the text. The principal result is of *qualitative* character. It consists in showing that for the same pressure drop the volume of fluid traversing a cross section of the pipe is greater for the steady flow than for turbulent motion. Or what amounts to the same thing the fall in pressure necessary to produce a given discharge from a pipe, is greater in the case of turbulent motion than for the steady flow. This result is in strict agreement with the experimental facts. In spite of its qualitative nature a certain interest may reasonably be attached to a result of this sort *first* since the demonstration is based on an exact treatment of the Navier-Stokes equations and equation of continuity and *second* since it is impossible to establish it if one confines himself to the usual exponential disturbances.

We proceed to a statement of the definition of stability, and the assumptions of continuity and differentiability underlying our discussion.

---

* Received July 29, 1941.

[1] *University of California Publications in Mathematics* (Los Angeles), vol. 1 (1942).

[2] In this connection see the comment by von Kármán, "The engineer grapples with non-linear problems," *Bulletin of the American Mathematical Society*, vol. 46 (1940), p. 664 (footnote).

2. Definition of stability. Continuity and differentiability assumptions. Let G denote a region completely filled by a viscous fluid and let us denote by $\bar{u}^{\alpha}$ $(\alpha = 1, 2, 3)$ the contravariant velocity components and by $\bar{p}$ the pressure of a stationary motion of the fluid in G. Similarly let u^{α} and p be the velocity components and pressure function of an arbitrary non-stationary fluid motion in G. The three quantities $\xi^{\alpha} = u^{\alpha} - \bar{u}^{\alpha}$ and $\xi^{4} = p - \bar{p}$ in any (dynamically) allowable coördinate system are the components of an invariant which we shall refer to as the *disturbance*[3] of the stationary motion.

The stationary motion will be said to be stable relative to a class D of disturbances in G if $\xi^{i} \to 0$ uniformly over G as $t \to \infty$ where the ξ^{i} $(i = 1, 2, 3, 4)$ are the components of an arbitrary disturbance of D.[4]

With regard to the characterization of the class D we must in the first place specify certain conditions of continuity and differentiability of the four components of the disturbances since the velocity components $\bar{u}^{\alpha}$ and the pressure $\bar{p}$ of the stationary motion and the corresponding functions u^{α} and p for those motions which give rise to the disturbances are assumed to satisfy the differential equation of continuity, and the Navier-Stokes equations as the basis of our discussion. Our requirements in this respect will be satisfied by the following two conditions which may be thought of as formulated relative to a rectangular cartesian coördinate system covering the region G.

(α). The components ξ^{i} $(i = 1, 2, 3, 4)$ are continuous functions of the space coördinates and the time in the region $\bar{G}$ (closure of G) and have continuous first and second partial derivatives with respect to the space coördinates and the time; in addition the fourth component $\xi^{4} = p - \bar{p}$ has continuous third partial derivatives with respect to the space coördinates.

(β). The components ξ^{i} $(i = 1, 2, 3, 4)$ of any disturbance and their partial derivatives in condition (α) are bounded uniformly (i. e., irrespective of the time) in $\bar{G}$.

In the above we have formulated the definition of stability and the fundamental assumptions of continuity and differentiability in a form sufficiently general to apply to a variety of problems on stability. In the following section we indicate specifically how this is to be understood in relation to the stability of fluids in pipes.

---

[3] The $\xi^{\alpha}(\alpha = 1, 2, 3)$ are the components of a contravariant vector and $\xi^{4} = p - \bar{p}$ is the component of a scalar under allowable coördinate transformations (which may involve the time t). Under the restricted class of transformations of the form $x^{\alpha} = x^{\alpha}(\bar{x})$, $t = \bar{t}$, where the x's are the spatial coördinates, the disturbance appears as a contravariant four-vector.

[4] By considering the condition $\xi^{i} \to 0$ as $t \to \infty$ rather than the more general condition $\xi^{i} \to$ const. as $t \to \infty$ we assume implicitly that the arbitrary constant which is involved in the pressure function p is so chosen that the const. in the convergence condition is equal to zero.

3. Equations of motion in cylindrical coördinates. In general coördinates x^1, x^2, x^3 the Navier-Stokes equations (in the absence of external forces) and the equation of continuity are

$$\mu\Delta u^\alpha = g^{\alpha\beta}\frac{\partial p}{\partial x^\beta} + \rho\frac{\partial u^\alpha}{\partial t} + \rho u^\beta u_{,\beta}{}^\alpha, \tag{3.1}$$

$$u_{,\alpha}{}^\alpha = 0, \text{ (equation of continuity)}. \tag{3.2}$$

All summations in these equations are over the range 1, 2, 3 and are made in accordance with the usual convention. The g's denote the components of the fundamental metric tensor and the comma indicates covariant differentiation. Also Δ is the Laplacian operator relative to the coördinate system employed. The quantity μ is the coefficient of viscosity and ρ denotes the density; we assume that both μ and ρ are constant throughout the motion of the fluid.

In our discussion of the flow of fluids in pipes we identify the coördinates x^1, x^2, x^3 with cylindrical coördinates r, ϕ, z in this order, the z-axis coinciding with the axis of the pipe $(r = 0)$. While the use of such coördinates may give rise to ambiguities in the determination of certain of the components of the velocity and their derivatives along the axis of the pipe it will be apparent that this causes no difficulty in the demonstrations.

In cylindrical coördinates the components of the metric tensor are

Covariant components: $g_{11} = g_{33} = 1,\ g_{22} = r^2,$
Contravariant components: $g^{11} = g^{33} = 1,\ g^{22} = 1/r^2.$

All other components of this tensor vanish. The non-vanishing components of the affine connection are $\Gamma_{22}^1 = -r,\ \Gamma_{12}^2 = \Gamma_{21}^2 = 1/r$. Using these the various quantities which occur in the above equations of motion can easily be calculated. For convenience, however, in carrying out certain of the operations indicated in the later sections we list the following formulas in which u, ω, w are used to stand for u^1, u^2, u^3 respectively.

$$\left\{\begin{aligned} &u_{,1}{}^1 = \frac{\partial u}{\partial r};\ u_{,2}{}^1 = \frac{\partial u}{\partial \phi} - r\omega;\ u_{,3}{}^1 = \frac{\partial u}{\partial z} \\ &u_{,1}{}^2 = \frac{\partial \omega}{\partial r} + \frac{\omega}{r};\ u_{,2}{}^2 = \frac{\partial \omega}{\partial \phi} + \frac{u}{r};\ u_{,3}{}^2 = \frac{\partial \omega}{\partial z} \\ &u_{,1}{}^3 = \frac{\partial w}{\partial r};\ u_{,2}{}^3 = \frac{\partial w}{\partial \phi};\ u_{,3}{}^3 = \frac{\partial w}{\partial z}\,. \end{aligned}\right. \tag{3.3}$$

$$(3.4)\qquad \begin{cases} u_{1,1} = \dfrac{\partial u}{\partial r}\,;\ u_{1,2} = \dfrac{\partial u}{\partial \phi} - r\omega\,;\ u_{1,3} = \dfrac{\partial u}{\partial z} \\[2mm] u_{2,1} = r^2 \dfrac{\partial \omega}{\partial r} + r\omega\,;\ u_{2,2} = r^2 \dfrac{\partial \omega}{\partial \phi} + ru\,;\ u_{2,3} = r^2 \dfrac{\partial w}{\partial z} \\[2mm] u_{3,1} = \dfrac{\partial w}{\partial r}\,;\ u_{3,2} = \dfrac{\partial w}{\partial \phi}\,;\ u_{3,3} = \dfrac{\partial w}{\partial z}\,. \end{cases}$$

$$(3.5)\qquad \begin{cases} \Delta u^1 = \dfrac{\partial^2 u}{\partial r^2} + \dfrac{1}{r^2}\dfrac{\partial^2 u}{\partial \phi^2} + \dfrac{\partial^2 u}{\partial z^2} + \dfrac{1}{r}\dfrac{\partial u}{\partial r} - \dfrac{2}{r}\dfrac{\partial \omega}{\partial \phi} - \dfrac{u}{r^2}\,, \\[2mm] \Delta u^2 = \dfrac{\partial^2 \omega}{\partial r^2} + \dfrac{1}{r^2}\dfrac{\partial^2 \omega}{\partial \phi^2} + \dfrac{\partial^2 \omega}{\partial z^2} + \dfrac{3}{r}\dfrac{\partial \omega}{\partial r} + \dfrac{2}{r^3}\dfrac{\partial u}{\partial \phi}\,, \\[2mm] \Delta u^3 = \dfrac{\partial^2 w}{\partial r^2} + \dfrac{1}{r^2}\dfrac{\partial^2 w}{\partial \phi^2} + \dfrac{\partial^2 w}{\partial z^2} + \dfrac{1}{r}\dfrac{\partial w}{\partial r}\,. \end{cases}$$

We emphasize, to avoid possible confusion, that in each of the three sets of formulas (3.3), (3.4) and (3.5) the u, ω, w denote the *contravariant* components of the velocity vector.

By a *steady flow* or Poiseuille motion in an infinite circular cylinder or pipe of radius a is meant the stationary motion for which the velocity components $\bar{u}^\alpha$ $(\alpha = 1, 2, 3)$ and pressure $\bar{p}$ are

$$(3.6)\qquad \bar{u}^1 = \bar{u}^2 = 0,\ \bar{u}^3 = w_0\left(1 - \frac{r^2}{a^2}\right),\ \bar{p} = -\frac{4\mu w_0}{a^2}\, z + \text{const.}$$

The constant w_0 is the velocity along the axis of the pipe and we suppose $w_0 \geqq 0$ so that the flow is in the positive direction of the z-axis.

A disturbance is said to have *spatial periodicity* if its components are periodic functions of period $l (> 0)$ in the direction of the axis of the pipe. In the case of spatial periodicity the cylinder is divided into *fixed* cells or blocks of length l (which may be different for different disturbances) such that the behavior of the disturbance in the pipe is completely determined by its behavior in any cell. The region G (interior) of the pipe can thus be replaced by any periodicity cell W in the discussion of the stability problem.

In the following we deal with disturbances in the pipe relative to the above stationary motion (3.6). We assume that the disturbances are spatially periodic. The condition of spatial periodicity of the components ξ^α $(\alpha = 1, 2, 3)$ of the disturbance appears as a mathematical convenience since it enables us to limit our attention to a particular periodicity cell. In the practical case of a (horizontal) pipe of finite length (one end of which is attached to a reservoir in which the level of the fluid is kept constant to secure a constant pressure at the entrance of the pipe) it is natural to associate the periodicity constant l in the theoretical discussion with the length of the pipe. The requirement of periodicity of the fourth component $\xi^4 = p - \bar{p}$ of the disturbance may then be interpreted to mean that the pressure difference between the ends of the pipe is the same for the disturbed motion as for the steady flow in conformity with the actual conditions of the experiment.

4. The energy equation. From the equations (3. 1) and (3. 2) and the corresponding equations for the steady flow (3. 6) we obtain the following differential equations of the disturbance

$$(4.1)\quad \begin{cases} \mu\Delta\xi^\alpha = g^{\alpha\beta}\dfrac{\partial(p-\bar{p})}{\partial x^\beta} + \rho\dfrac{\partial\xi^\alpha}{\partial t} + \rho\xi^\beta\xi_{,\beta}{}^\alpha + \rho\bar{u}^\beta\xi_{,\beta}{}^\alpha + \rho\xi^\beta\bar{u}_{,\beta}{}^\alpha, \\ \xi_{,\alpha}{}^\alpha \equiv \dfrac{\partial\xi^1}{\partial r} + \dfrac{\xi^1}{r} + \dfrac{\partial\xi^2}{\partial\phi} + \dfrac{\partial\xi^3}{\partial z} = 0. \end{cases}$$

From these equations we deduce

$$(4.2)\quad \begin{aligned} \mu\Delta(g_{\alpha\beta}\xi^\alpha\xi^\beta) = 2\mu g_{\alpha\beta}g^{\gamma\delta}\xi_{,\gamma}{}^\alpha\xi_{,\delta}{}^\beta + 2[(p-\bar{p})\xi^\alpha]_{,\alpha} + \rho\frac{\partial}{\partial t}(g_{\alpha\beta}\xi^\alpha\xi^\beta) \\ + \rho[(g_{\alpha\beta}\xi^\alpha\xi^\beta)\xi^\gamma]_{,\gamma} + \rho[(g_{\alpha\beta}\xi^\alpha\xi^\beta)\bar{u}^\gamma]_{,\gamma} + 2\rho D_{\alpha\beta}\xi^\alpha\xi^\beta, \end{aligned}$$

where

$$D_{\alpha\beta} = \tfrac{1}{2}(\bar{u}_{\alpha,\beta} + \bar{u}_{\beta,\alpha})$$

are the components of the *deformation tensor* of the steady flow. Using (3. 6) we find $D_{13} = D_{31} = -w_0 r/a^2$. All other components of the deformation tensor vanish.

Now integrate both members of (4. 2) over the periodicity cell W. Applying Green's theorem with the condition $\xi^\alpha = 0$ over the surface of the pipe and taking account of the property of periodicity of the disturbance we obtain the *energy equation*

$$(4.3)\quad \frac{dK}{dt} + \mu\int_W g_{\alpha\beta}g^{\gamma\delta}\xi_{,\gamma}{}^\alpha\xi_{,\gamma}{}^\alpha\xi_{,\delta}{}^\beta dV + \rho\int_W D_{\alpha\beta}\xi^\alpha\xi^\beta dV = 0,$$

where

$$K = \tfrac{1}{2}\rho\int_W g_{\alpha\beta}\xi^\alpha\xi^\beta dV$$

is the kinetic energy of the disturbance in W.

5. Modified energy equation. Consider

$$(5.1)\quad \int_W \xi^3 dV = \int_W \frac{\partial}{\partial z}(z\xi^3)\,dV - \int_W z\frac{\partial\xi^3}{\partial z}\,dV = l[Q(t) - q],$$

where $Q(t)$ denotes the volume of fluid which traverses a cross section of pipe in unit time for the disturbed motion and q denotes the corresponding quantity for the steady flow (3. 6). In fact the second integral in the middle number of this equation is easily seen to vanish by the equation of continuity and the first integral in the middle member immediately yields the right member of the equation.

Now take $\alpha = 3$ in the first equation of (4.1), multiply both members of the equation by the scalar quantity $1 - r^2/a^2$ and then integrate over the periodicity cell W. The integrals of all terms of the equation are found to vanish with the exception of the following:

$$\mu \int_W \left(1 - \frac{r^2}{a^2}\right) g^{\alpha\beta} \xi_{,\alpha\beta}{}^3 dV = -\frac{4l\mu}{a^2} [Q(t) - q],$$

$$\rho \int_W \left(1 - \frac{r^2}{a^2}\right) \frac{\partial \xi^3}{\partial t} dV = \rho \frac{d}{dt} \int_W \left(1 - \frac{r^2}{a^2}\right) \xi^3 dV,$$

$$\rho \int_W \left(1 - \frac{r^2}{a^2}\right) \xi^\alpha \xi_{,\alpha}{}^3 dV = \frac{2\rho}{a^2} \int_W r \xi^1 \xi^3 dV.$$

In the derivation of these formulas and in the proof that the remaining integrals in question are equal to zero we have made use of (5.1), the equation of continuity, Green's theorem and the property of spatial periodicity of the disturbance as in the derivation of (4.3). This leads to the following equation:

$$(5.2) \qquad w_0 \rho \frac{d}{dt} \int_W \left(1 - \frac{r^2}{a^2}\right) \xi^3 dV + \frac{4 w_0 \mu l}{a^2} [Q(t) - q] + \frac{2 w_0 \rho}{a^2} \int_W r \xi^1 \xi^3 dV = 0.$$

Now when we substitute the explicit values of the components of the deformation tensor into the last term of (4.3) we find that this term is the negative of the last term of (5.2). Hence adding corresponding members of (4.3) and (5.2) we obtain

$$(5.3) \qquad \frac{dF}{dt} + \mu \int_W g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}{}^\alpha \xi_{,\delta}{}^\beta dV + \frac{4 w_0 \mu l}{a^2} [Q(t) - q] = 0,$$

where

$$F = K + w_0 \rho \int_W \left(1 - \frac{r^2}{a^2}\right) \xi^3 dV.$$

We shall refer to (5.3) as the *modified energy equation.* It is easily seen that the function F in this equation is the difference between the kinetic energies of the disturbed motion and the steady flow (3.6) in the periodicity cell W.

6. The stability theorem. In this section we state the main results of the paper. The proofs will be given in the two following sections:

STABILITY THEOREM. *Steady flow in a pipe is stable relative to the class of spatially periodic disturbances satisfying the conditions* (α) *and* (β) *in Section* 2 *and such that* $Q(t) \geqq q$ *at all times* t.

The meaning of the condition $Q(t) \geqq q$ is that the discharge (volume of fluid traversing a cross section of the pipe per unit time) is never less for the disturbed motion than for the steady flow. It is an open question if there exist spatially periodic disturbances (other than the zero disturbance) satisfying the condition $Q(t) \geqq q$ and we have made no attempt to treat this question (which would require an essentially different analysis) in the present paper. The above stability result thus means effectively that *if* any motion exists which gives rise to such a disturbance it converges uniformly to the steady flow. Let us say (1) that a motion is *established* if the discharge is constant and (2) that a motion is *turbulent* if the corresponding disturbance ξ does not approach zero uniformly as $t \rightarrow \infty$. In the case of *established turbulence* the above stability theorem then leads to the result that the discharge is such that $Q < q$.

Let us now make the substitution $\xi^a = u^a - \bar{u}^a$ in the integral which appears explicitly in (5.3). The resulting equation can then be reduced to the form

$$\frac{dF}{dt} + \mu \int_W g_{\alpha\beta} g^{\gamma\delta} u_{,\gamma}^{\alpha} u_{,\delta}^{\beta} dV - \frac{4 w_0 \mu l}{a^2} Q(t) = 0. \tag{6.1}$$

If the last term in this equation is essentially positive, i. e., $Q(t) \leqq 0$, the discussion in Section 7 (with ξ^a replaced by u^a) leads to the conclusion that $F(t) \rightarrow -\infty$ as $t \rightarrow \infty$. This is in contradiction with the condition (β) in Section 2. For established turbulence it therefore follows that $0 < Q < q$, i. e., *in the case of established turbulence the discharge is positive and less than the discharge for the steady flow.*

This latter result can also be stated in the following terms: *the pressure drop required to produce a given discharge is greater in the case of established turbulence than for a steady flow in the pipe.* This result is in strict agreement with observation and naturally expresses a fact of considerable practical importance.

7. First part of convergence proof. We shall now prove that *the first three components* ξ^a $(\alpha = 1, 2, 3)$ *of any disturbance approach zero uniformly over* $\bar{W}$ *(closure of* W*) as* $t \rightarrow \infty$. The assumptions underlying this proof are those appearing in the Stability Theorem in Section **6**.

Suppose that ξ^a does not approach zero uniformly over $\bar{W}$ as $t \rightarrow \infty$. Then $g_{\alpha\beta}\xi^\alpha\xi^\beta$ does not. Hence we can find some $\epsilon > 0$, a time sequence $t_1, t_2, t_3, \cdots$ with $t_i > t_j$ $(i > j)$ and $t_n \rightarrow \infty$ as $n \rightarrow \infty$, and a point sequence $P_1, P_2, P_3, \cdots$ with $P_n \subset \bar{W}$ such that

$$(g_{\alpha\beta}\xi^\alpha\xi^\beta)_{P_n} > \epsilon, \text{ for } t = t_n \qquad (n = 1, 2, 3, \cdots). \tag{7.1}$$

We can suppose, in particular, that the time sequence is so chosen that $t_{n+1} > t_n + 1$ for $n = 1, 2, 3, \cdots$ as is evidently possible. It is convenient to have some such inequality at a certain place in the following discussion.

We now show on the basis of (7.1) that *the non-negative quantities*

$$\int_W (g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}{}^{\alpha} \xi_{,\delta}{}^{\beta})_{t_n} dV, \qquad (n = 1, 2, 3, \cdots), \tag{7.2}$$

are bounded away from zero. First, for no value of n will the integral (7.2) be equal to zero. For if such were the case this would imply $\xi_{,\beta}{}^{\alpha} = 0$ for $t = t_n$ and since $\xi^{\alpha} = 0$ on the surface of the pipe it would follow that $\xi^{\alpha} = 0$ over $\bar{W}$ at the time t_n. This is in contradiction with (7.1). If the quantities (7.2) are *not* bounded away from zero we can find a subsequence of t_n (which we continue to denote by t_n for simplicity of notation) such that

$$\int_W (g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}{}^{\alpha} \xi_{,\delta}{}^{\beta})_{t_n} dV \to 0, \text{ as } n \to \infty. \tag{7.3}$$

Now let x^1, x^2, x^3 be the coördinates of a rectangular coördinate system in which conditions (α) and (β) in Section 2 strictly apply by hypothesis. Denote the quantities ξ^{α} by η^{α} with respect to this coördinate system. Then by the condition (β) we can find a positive constant M such that

$$|\eta^{\alpha}(x, t_n)|, \left|\frac{\partial \eta^{\alpha}(x, t_n)}{\partial x^{\beta}}\right|, \left|\frac{\partial^2 \eta^{\alpha}(x, t_n)}{\partial x^{\beta} \partial x^{\gamma}}\right| < M \tag{7.4}$$

over $\bar{W}$ for $n = 1, 2, 3, \cdots$. From the first of the inequalities (7.4) the sequence $\eta^{\alpha}(x, t_n)$ is uniformly bounded in absolute value in $\bar{W}$ and from the second inequality the sequence is *equi-continuous* in $\bar{W}$. Since $\bar{W}$ is a closed and bounded region it follows by the *Theorem of Ascoli*[5] that this sequence contains a subsequence $\eta^{\alpha}(x, t_{n_k})$ which converges uniformly to a continuous function $\zeta^{\alpha}(x)$ in $\bar{W}$. Similarly from the second and third inequalities (7.4) the sequence $\partial \eta^{\alpha}(x, t_{n_k})/\partial x^{\beta}$, $(k = 1, 2, 3, \cdots)$, contains a subsequence which converges uniformly to a continuous function $\zeta_{\beta}{}^{\alpha}(x)$ in $\bar{W}$. Using $n = 1, 2, 3, \cdots$ to denote the terms of this latter sequence for simplicity of notation we thus have

$$\eta^{\alpha}(x, t_n) \to \zeta^{\alpha}(x), \frac{\partial \eta^{\alpha}(x, t_n)}{\partial x^{\beta}} \to \zeta_{\beta}{}^{\alpha}(x) \tag{7.5}$$

uniformly in $\bar{W}$ as $n \to \infty$ where $\zeta^{\alpha}(x)$ and $\zeta_{\beta}{}^{\alpha}(x)$ are continuous functions of the coördinates x^1, x^2, x^3 of $\bar{W}$. Finally it follows from (7.5) and a well known theorem on uniform convergence that the functions $\zeta^{\alpha}(x)$ have continuous first partial derivatives and in fact that $\partial \zeta^{\alpha}(x)/\partial x^{\beta} = \zeta_{\beta}{}^{\alpha}(x)$ in $\bar{W}$.

---

[5] See O. D. Kellogg, *Foundations of Potential Theory*, Berlin (1929), p. 265.

Using the condition (7.3) and the fact that the convergence in (7.5) is uniform we now have

$$\lim_{n\to\infty}\int_W (g_{\alpha\beta}g^{\gamma\delta}\xi_{,\gamma}{}^{\alpha}\xi_{,\delta}{}^{\beta})_{t_n}dV = \int_W \sum_{\alpha,\beta} \frac{\partial\zeta^\alpha}{\partial x^\beta}\frac{\partial\zeta^\alpha}{\partial x^\beta}\,dV = 0.$$

It follows that $\partial\zeta^\alpha/\partial x^\beta = 0$ and hence $\zeta^\alpha = 0$ over $\bar{W}$ since $\zeta^\alpha = 0$ over the surface of the pipe by the boundary condition. Hence $\eta^\alpha(t_n) \to 0$ and hence $\xi^\alpha(t_n) \to 0$ uniformly over $\bar{W}$ as $n \to \infty$ and this contradicts (7.1). It follows that the integrals (7.2) are bounded away from zero as above stated.

Now consider the *graph* of the function

$$y(t) = \mu\int_W g_{\alpha\beta}g^{\gamma\delta}\xi_{,\gamma}{}^{\alpha}\xi_{,\delta}{}^{\beta}dV + \frac{4w_0\mu l}{a^2}\,[Q(t) - q] \tag{7.6}$$

with t measured along the horizontal axis and y along the vertical axis of a rectangular cartesian coördinate system. Since the second term in the right member of (7.6) is non-negative by the hypothesis of the Stability Theorem it follows from the result just proved that we can find a number $k > 0$ such that $y(t_n) > k$ for $n = 1, 2, 3 \cdots$. It also follows from condition (β) in Section 2 that we can find a positive number N such that $|\,dy/dt\,| < N$ at all times t.

Now let δ be a positive number and consider the intervals $t_n - \delta \leqq t \leqq t_n + \delta$ for $n = 1, 2, 3, \cdots$. We take δ so small that the successive intervals have no point in common; this is possible in view of the condition $t_{n+1} > t_n + 1$ imposed at the beginning of the section. Let us furthermore take δ so small that $A = 2k\delta - N\delta^2 > 0$. Now if A_n denotes the area underneath the above graph and above the interval $t_n - \delta \leqq t \leqq t_n + \delta$ it follows from the condition $|\,dy/dt\,| < N$ that $A_n > A$ for $n = 1, 2, 3, \cdots$. Hence by integration of the modified energy equation (5.3) and the fact that $y(t) \geqq 0$ for all t we obtain

$$F(t_n + \delta) - F(t_1 - \delta) = -\int_{t_1-\delta}^{t_n+\delta} y(t)\,dt < -\sum_{a=1}^{n} A_a < -nA.$$

It follows that $F(t) \to -\infty$ as $t \to \infty$. But this contradicts the hypothesis that the components ξ^α of the disturbance are uniformly bounded and thus completes the proof of the italicized statement at the beginning of the section.

8. Second part of convergence proof. It remains to prove the uniform convergence of the fourth component $\xi = p - \bar{p}$ of the disturbance. For this purpose we first multiply both members of the first equation (4.1) by $\partial\xi^4/\partial x^\alpha$ and sum on the repeated index. We then integrate both members of the resulting equation over W. This gives rise to the following integrals

$$(8.1)\qquad \int_W \xi^4_{,\alpha}\Delta\xi^\alpha dV = \int_W [\xi^4_{,\alpha} g^{\gamma\delta}\xi^\alpha_{,\gamma}]_{,\delta} dV - \int_W [\xi^4_{,\alpha\beta} g^{\beta\gamma}\xi^\alpha]_{,\gamma} dV + \int_W \xi^4_{,\alpha\beta\gamma} g^{\beta\gamma}\xi^\alpha dV,$$

$$(8.2)\qquad \int_W \xi^4_{,\alpha}\frac{\partial\xi^\alpha}{\partial t}\, dV = \frac{d}{dt}\int_W [\xi^4\xi^\alpha]_{,\alpha} dV - \int_W \frac{\partial^2\xi^4}{\partial x^\alpha\partial t}\,\xi^\alpha dV,$$

$$(8.3)\qquad \int_W \xi^4_{,\alpha}\xi^\beta\xi^\alpha_{,\beta} dV = \int_W [\xi^4_{,\alpha}\xi^\alpha\xi^\beta]_{,\beta} dV - \int_W \xi^4_{,\alpha\beta}\xi^\alpha\xi^\beta dV,$$

$$(8.4)\qquad \int_W \xi^4_{,\alpha}\bar{u}^\beta\xi^\alpha_{,\beta} dV = \int_W [\xi^4_{,\alpha}\bar{u}^\beta\xi^\alpha]_{,\beta} dV - \int_W \xi^4_{,\alpha\beta}\bar{u}^\beta\xi^\alpha dV,$$

$$(8.5)\qquad \int_W \xi^4_{,\alpha}\xi^\beta\bar{u}^\alpha_{,\beta} dV = \int_W [\xi^4_{,\alpha}\xi^\beta\bar{u}^\alpha]_{,\beta} dV - \int_W \xi^4_{,\alpha\beta}\xi^\alpha\bar{u}^\beta dV,$$

in which ξ^4 is treated as a scalar. By Green's theorem the first integral in the right member of (8.1) becomes

$$(8.6)\qquad -\int_B \xi^4_{,\alpha}\xi^\alpha_{,\beta} n^\beta dS = -\int_B \xi^4_{,\alpha}\frac{d\xi^\alpha}{dn}\, dS - \int_B \xi^4_{,\alpha}\xi^\beta\Gamma^\alpha_{\beta\gamma} n^\gamma dS,$$

over the surface or outer boundary B of the periodicity cell W. Here n denotes the inward unit normal to B and $d\xi^\alpha/dn$ is the ordinary derivative of ξ^α along this normal. Since the ξ^α approach zero uniformly over $\bar{W}$ by the result of Section **7** it follows readily (for example, by condition (β) and application of the mean value theorem) that $d\xi^\alpha/dn \to 0$ uniformly over B as $t \to \infty$. Hence the first integral in the right member of (8.6) approaches zero as $t \to \infty$. The second integral in the right member of (8.6) vanishes by the boundary condition ($\xi^\alpha = 0$ on B). Similarly all remaining integrals in the right members of (8.1), $\cdots$, (8.5) either (1) vanish by Green's Theorem and the assumption of spatial periodicity or (2) approach zero at $t \to \infty$ in consequence of condition (β) and the result of Section **7**. Hence the integrals in the left members of these equations approach zero as $t \to \infty$. The above procedure of multiplying (4.1) by $\partial\xi^4/\partial x^\alpha$ and integrating over W therefore leads to the result that

$$(8.7)\qquad \int_W g^{\alpha\beta}\frac{\partial(p-\bar{p})}{\partial x^\alpha}\,\frac{\partial(p-\bar{p})}{\partial x^\beta}\, dV \to 0, \text{ as } t \to \infty.$$

Let us now interpret the coördinates x^1, x^2, x^3 as rectangular cartesian coördinates (transformation to a rectangular system) as was done in Section **7**. Putting $v_\alpha = \partial(p-\bar{p})/\partial x^\alpha$ for simplicity the condition (8.7) then becomes

$$(8.8)\qquad L(t) \equiv \int_W \Sigma v_\alpha v_\alpha dV \to 0, \text{ as } t \to \infty.$$

We shall now show that

$$G(x,t) \equiv \Sigma\, v_\alpha(x,t) v_\alpha(x,t) \rightarrow 0, \text{ as } t \rightarrow \infty, \tag{8.9}$$

uniformly over $\bar{W}$ in consequence of (8.8) and condition (β) in Section 2. Since $G(x,t)$ is a continuous function of x^1, x^2, x^3 in $\bar{W}$ (closed and bounded region) at any time t, this function will assume its maximum value $G(t)$ at some point of $\bar{W}$. The condition (8.9) implies

$$G(t) \rightarrow 0, \text{ as } t \rightarrow \infty, \tag{8.10}$$

and conversely (8.10) implies (8.9). Hence it suffices to prove (8.10).

Proof of (8.10). By condition (β) the one parameter family of functions $v_\alpha(x,t)$ is *equi-continuous* in $\bar{W}$ (the parameter being t). Hence the family $G(x,t)$ is equi-continuous in $\bar{W}$. Explicitly this means that if any $\epsilon > 0$ is chosen we can find a $\delta > 0$ such that

$$| G(x',t) - G(x'',t) | < \epsilon, \text{ for } | x' - x'' | < \delta, \tag{8.11}$$

and for all times t, where we have used $| x' - x'' |$ to denote the (Euclidean) distance between the points x' and x'' in $\bar{W}$.

Assume that (8.10) is not satisfied. Then we can find some constant $\eta > 0$ such that for any time t_0 we shall have $G(t_1) \geqq \eta$ for some $t_1 \geqq t_0$. Hence there will exist a time sequence τ_n with $\tau_n \rightarrow \infty$ as $n \rightarrow \infty$ such that $G(\tau_n) \geqq \eta$ for $n = 1, 2, 3, \cdots$. Let P_n be a point of $\bar{W}$ at which $G(x, \tau_n)$ assumes its maximum value $G(\tau_n)$. Since $\bar{W}$ is closed and bounded the sequence $P_1, P_2, P_3, \cdots$ will admit a subsequence (for simplicity of notation we use P_n also to denote this subsequence) which will approach a limit point P in $\bar{W}$, i.e. $P_n \rightarrow P$ as $n \rightarrow \infty$.

Assume *first* that P is a point of W (interior point). Let λ be the interior of a sphere having center at P and radius r so small that its closure $\bar{\lambda}$ is contained in W. Take $\epsilon = \eta/2$ in (8.11). Then chose r so small that if x' and x'' are any two points in λ the condition $| x' - x'' | < \delta$ in (8.11) will be satisfied. For $n \geqq N$ (sufficiently large) we will have P_n in λ. Since $G(P_n, \tau_n) = G(\tau_n)$ the equi-continuity condition (8.11) then gives $| G(\tau_n) - G(x, \tau_n) | < \eta/2$, $n \geqq N$, for all points x in λ. Hence for x in λ we have

$$G(x, \tau_n) > G(\tau_n) - \eta/2 \geqq \eta/2, \quad n \geqq N, \tag{8.12}$$

when use is made of the above assumption that $G(\tau_n) \geqq \eta$ for $n = 1, 2, 3, \cdots$. From (8.12) we now obtain

$$L(\tau_n) = \int_W G(x, \tau_n)\,dV \geqq \int_\lambda G(x, \tau_n)\,dV \geqq \frac{2}{3}\pi r^3 \eta$$

for $n \geqq N$. But this contradicts (8.8) and so the above assumption is false. Hence (8.10), and hence (8.9), are satisfied.

Assume in the *second* place that P is a boundary point of W. Consider a sphere with center at P and radius r and let λ be the intersection of this sphere with W. Choose r so small that the δ condition of the first assumption will be satisfied. The above discussion then applies and we are led to the inequalities $L(\tau_n) \geqq \pi r^3 k$ for $n \geqq N$ where k is a suitable positive constant less than unity. Thus again we have a contradiction with (8.8). This completes the proof of (8.9) and (8.10).

From (8.9) or (8.10) we readily see that the quantities $v_\alpha(x, t) \to 0$ uniformly over $\bar{W}$ as $t \to \infty$. Since the pressure function is determined at most to within an additive constant by the equations of motion we can suppose that the pressures p and $\bar{p}$ are always the same at a fixed point T in W. Then the above convergence condition for the derivatives of $p - \bar{p}$ leads to the result that $p - \bar{p} \to 0$ uniformly over $\bar{W}$ as $t \to \infty$. This completes the proof of the Stability Theorem of Section **6**.

9. Extension of the stability theorem.

Consider the equation

$$\frac{dJ}{dt} + \mu \int_W g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}^{\alpha} \xi_{,\delta}^{\beta}\,dV + \epsilon(t) = 0, \tag{9.1}$$

where J is the integral over W of a polynomial in the ξ's and their first and second derivatives and $\epsilon(t) \to 0$ as $t \to \infty$. The discussion in Sections **7** and **8** shows that *the steady flow in a pipe is stable relative to the class of spatially periodic disturbances satisfying the conditions* (α) *and* (β) *in Section* **2** *and the differential equation* (9.1). As in Section **7** we show that the integrals (7.2) are bounded away from zero. We then consider the equation (7.6) in which the last term is replaced by $\epsilon(t)$. Since $\epsilon(t) \to 0$ as $t \to \infty$ it follows that $y(t_n) > k > 0$ provided that $n \geqq N$ (sufficiently large). Subject to this restriction the remainder of the first part of the convergence proof follows as in Section **7** and leads to the contradiction that $J(t) \to -\infty$ as $t \to \infty$. The second part of the convergence proof follows as in Section **8** without modification.

10. Boundary conditions and stability.

Let us now take $\alpha = 3$ in the first equation (4.1) and then integrate both members of the equation over the periodicity cell W. After some reduction the following equation is obtained

$$l \frac{dQ}{dt} - a\nu \int_B \frac{\partial \xi^3}{\partial r} \, dS = 0, \tag{10.1}$$

where B denotes the surface or outer boundary of the periodicity cell and $\nu(=\mu/\rho)$ is the kinematical viscosity; the remaining quantities in (10.1) have already been defined.

The discussion in Section **6** shows the dependence of the stability of the steady flow on the discharge $Q(t)$. The dependence of $Q(t)$ in turn on the state of motion of the fluid at the boundary is shown in (10.1). If the boundary integral in this equation vanishes for all values of the time t we have $Q(t) = \text{const.}$ which is the condition for established motion as defined in Section **6**. Now the integral

$$\int_B \frac{\partial \xi^3}{\partial r} \, dS \tag{10.2}$$

cannot be bounded away from zero for it would then follow from (10.1) that $Q(t) \to \pm \infty$ as $t \to \infty$ thereby contradicting the condition (β) in Section **2**. If this integral oscillates between positive and negative values there will be a continual increase and decrease in the discharge $Q(t)$. Actually a phenomenon of this type is observed to occur experimentally near a critical value of the Reynolds number R. The determination of the conditions underlying such an oscillatory behavior of the integral are thus apparently connected with the problem of determining a critical value of R and the demands of this problem exceed the qualitative analysis of the present paper.

11. Stability theorem for exponential disturbances. Attempts have been made to treat the problem of the stability of steady flow in pipes by considering spatially periodic disturbances of the type[6]

$$\xi^k = R[f^k(r, \phi) e^{\sigma t + i\lambda x}], \qquad (k = 1, 2, 3, 4), \tag{11.1}$$

where R denotes the real part of the quantities in the brackets. Here the functions f^k may be complex and σ may be a complex constant. Also λ is a real constant and can be supposed to be positive without loss of generality. The periodicity l is then given by $l = 2\pi/\lambda$. It is sufficient to suppose that the functions $f^k(r, \phi)$ are continuous and have continuous first and second partial derivatives.

---

[6] An excellent comprehensive account of these and other investigations has been given by J. L. Synge, "Hydrodynamical Stability," *American Mathematical Society Semicentennial Publications* II: Addresses, 1938, pp. 227-269. Synge deals specifically with disturbances of this type on p. 250 and again on p. 264 such disturbances are presumably at the basis of his discussion.

It follows immediately from (11.1) that $F = K$ and $Q(t) = q$. Hence (5.3) reduces to

$$\frac{dK}{dt} + \mu \int_W g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}{}^{\alpha} \xi_{,\delta}{}^{\beta} dV = 0. \tag{11.2}$$

Now put $\sigma = \sigma_1 + i\sigma_2$ and $f^k = f_1{}^k + if_2{}^k$ where σ_1, σ_2, $f_1{}^k$ and $f_2{}^k$ are real. Substituting into (11.1) we find

$$\xi^k = e^{\sigma_1 t}[\cos \sigma_2 t (f_1{}^k \cos \lambda z - f_2{}^k \sin \lambda z) - \sin \sigma_2 t (f_1{}^k \sin \lambda z + f_2{}^k \cos \lambda z)].$$

Using these expressions for the ξ^k we may deduce the following relations

$$\int_W (\xi^k)^2 dV = \frac{2\pi}{\lambda} e^{2\sigma_1 t} \int_\Omega [(f_1{}^k)^2 + (f_2{}^k)^2] dS,$$

where Ω denotes any right cross section of the pipe. Hence the integral for the kinetic energy K (Section **4**) becomes

$$K = \frac{\pi\rho}{\lambda} e^{2\sigma_1 t} \int_\Omega g_{\alpha\beta}(f_1{}^\alpha f_1{}^\beta + f_2{}^\alpha f_2{}^\beta) dS.$$

Finally substituting this expression for K into (11.2) we have

$$\frac{2\pi\rho\sigma_1}{\lambda} e^{2\sigma_1 t} \int_\Omega g_{\alpha\beta}(f_1{}^\alpha f_1{}^\beta + f_2{}^\alpha f_2{}^\beta) dS + \mu \int_W g_{\alpha\beta} g^{\gamma\delta} \xi_{,\gamma}{}^{\alpha} \xi_{,\delta}{}^{\beta} dV = 0. \tag{11.3}$$

If $\sigma_1 \geqq 0$ each of the terms in (11.3) is non-negative and hence must vanish in consequence of this equation. The vanishing of the second term in (11.3) implies that $\xi_{,\beta}{}^{\alpha} = 0$ and when these conditions are combined with the boundary condition ($\xi^\alpha = 0$ on the surface of the pipe) it follows that $\xi^\alpha = 0$ ($\alpha = 1, 2, 3$) over the periodicity cell. Hence from the Navier-Stokes equation the space derivatives of ξ^4 must vanish; from this and the above expression for ξ^4 we easily see that $\xi^4 = 0$. We thus arrive at the case of a *zero disturbance*. Hence if ξ is not a zero disturbance it follows from (11.3) that $\sigma_1 < 0$. This gives us the following stability theorem: *The steady flow in a pipe is stable relative to the class of disturbances represented by* (11.1).

The above disturbances (11.1) are a special case of disturbances represented by

$$f_1(r, \phi, t) e^{2\pi iz/l} + f_2(r, \phi, t) e^{4\pi iz/l} + \cdots$$

which are spatially periodic of period l in the direction of the z-axis. For such disturbances the equation (5.3) likewise reduces to (11.2). Hence the Stability Theorem in Section **6** applies, i. e. the steady flow is stable relative to this class of disturbances under the conditions on the functions $f_1(r, \phi, t)$, $f_2(r, \phi, t)$, $\cdots$ implied by the assumptions of the Stability Theorem.

UNIVERSITY OF CALIFORNIA
AT LOS ANGELES.

NOTE ON THE EXPRESSIONS FOR THE STRAINS IN A BENT, THIN SHELL.*

By ERIC REISSNER.

In a recent paper [5] the linear theory of bending and stretching of thin elastic shells has been discussed anew. The purpose of the present note is to make that discussion more complete by deriving from the general formulas of the preceding paper expressions for the components of strain and for the stress resultant strain relations which are exact up to terms of that order of magnitude which is relevant in view of the basic assumption of shell theory that the normal to the undeformed middle surface is deformed into the normal to the deformed middle surface (Navier's hypothesis).

The results in [5] which form the starting point of the present discusion [1] are equivalent to the following formulas for the significant components of strain of the bent shell,

$$(1)\qquad \epsilon_1 = \frac{(\partial/\partial\xi_1)(u_1+\zeta u_1')}{\alpha_1(1+\zeta/R_1)} + \frac{(u_2+\zeta u_2')(\partial/\partial\xi_2)(\alpha_1(1+\zeta/R_1))}{\alpha_1\alpha_2(1+\zeta/R_1)(1+\zeta/R_2)} + \frac{w/R_1}{(1+\zeta/R_1)}$$

a corresponding expression for ϵ_2 and

$$(2)\qquad \gamma_{12} = \frac{\alpha_2(1+\zeta/R_2)}{\alpha_1(1+\zeta/R_1)}\frac{\partial}{\partial\xi_1}\left(\frac{u_2+\zeta u_2'}{\alpha_2(1+\zeta/R_2)}\right) + \frac{\alpha_1(1+\zeta/R_1)}{\alpha_2(1+\zeta/R_2)}\frac{\partial}{\partial\xi_2}\left(\frac{u_1+\zeta u_1'}{\alpha_1(1+\zeta/R_1)}\right).$$

In the preceding discussion it had been assumed, in conformity with A. E. H. Love [3] that the terms in (1) and (2) having ζ/R_n as a factor are negligible and on the basis of this assumption expressions of the following sort were obtained

$$(3)^2\qquad \epsilon_1 \approx \epsilon_1^0 + \zeta\kappa_1, \qquad \epsilon_2 \approx \epsilon_2^0 + \zeta\kappa_2, \qquad \gamma_{12} \approx \gamma_{12}^0 + \zeta\tau.$$

It seems, however, that in order to be consistent *all* terms in (1) and (2) which are linear in ζ should be retained (and none which contain higher powers of ζ since such terms are already neglected by working with Navier's hypothesis). If this is done there follows

---

* Received August 22, 1941.

[1] The symbols used in this note without explanation are defined in [5].

[2] See equations (36) to (39) in [5].

(4) $$\epsilon_1 = \epsilon_1{}^0 + \zeta\kappa_1{}^*, \qquad \epsilon_2 = \epsilon_2{}^0 + \zeta\kappa_2{}^*, \qquad \gamma_{12} = \gamma_{12}{}^0 + \zeta\tau^*$$

where

(5) $$\kappa_1{}^* = \kappa_1 - \frac{\epsilon_1{}^0}{R_1} - \frac{u_2}{\alpha_1\alpha_2}\left[\frac{1}{R_2}\frac{\partial\alpha_1}{\partial\xi_2} - \frac{\partial}{\partial\xi_2}\left(\frac{\alpha_1}{R_1}\right)\right]$$

(6) $$\kappa_2{}^* = \kappa_2 - \frac{\epsilon_2{}^0}{R_2} - \frac{u_1}{\alpha_1\alpha_2}\left[\frac{1}{R_1}\frac{\partial\alpha_2}{\partial\xi_1} - \frac{\partial}{\partial\xi_1}\left(\frac{\alpha_2}{R_2}\right)\right]$$

(7) $$\tau^* = \tau - \left(\frac{1}{R_1} - \frac{1}{R_2}\right)\omega - \left[\frac{\alpha_2}{\alpha_1}\frac{\partial}{\partial\xi_1}\left(\frac{u_2}{\alpha_2 R_2}\right) + \frac{\alpha_1}{\alpha_2}\frac{\partial}{\partial\xi_2}\left(\frac{u_1}{\alpha_1 R_1}\right)\right]$$

and

(8) $$\omega = \frac{\alpha_2}{\alpha_1}\frac{\partial}{\partial\xi_1}\left(\frac{u_2}{\alpha_2}\right) - \frac{\alpha_1}{\alpha_2}\frac{\partial}{\partial\xi_2}\left(\frac{u_1}{\alpha_1}\right).$$

If we use these modified expressions for the strain components we have, again in order to be consistent, to retain in the expressions for the stress resultants[3] all terms which contain third or lower powers of the shell thickness h. This leads to the following expressions

(9) $$N_{11} = \frac{E}{1-\nu^2}\{h(\epsilon_1{}^0 + \nu\epsilon_2{}^0) + \frac{h^3}{12R_2}(\kappa_1{}^* + \nu\kappa_2{}^*)\}$$

(10) $$N_{12} = G\{h\gamma_{12}{}^0 + \frac{h^3}{12R_2}\tau^*\}$$

(11) $$N_{21} = G\{h\gamma_{12}{}^0 + \frac{h^3}{12R_1}\tau^*\}$$

(12) $$N_{22} = \frac{E}{1-\nu^2}\{h(\epsilon_2{}^0 + \nu\epsilon_1{}^0) + \frac{h^3}{12R_1}(\kappa_2{}^* + \nu\kappa_1{}^*)\}$$

(13) $$M_{11} = \frac{Eh^3}{12(1-\nu^2)}\{(\kappa_1{}^* + \nu\kappa_2{}^*) + \frac{1}{R_2}(\epsilon_1{}^0 + \nu\epsilon_2{}^0)\}$$

(14) $$M_{12} = \frac{Gh^3}{12}\{\tau^* + \frac{1}{R_2}\gamma_{12}{}^0\}$$

(15) $$M_{21} = \frac{Gh^3}{12}\{\tau^* + \frac{1}{R_1}\gamma_{12}{}^0\}$$

(16) $$M_{22} = \frac{Eh^3}{12(1-\nu^2)}\{(\kappa_2{}^* + \nu\kappa_1{}^*) + \frac{1}{R_1}(\epsilon_2{}^0 + \nu\epsilon_1{}^0)\}.$$

It is to be noted that if the power series development of the strain components (1) and (2) were continued so as to include terms in ζ^2, these terms would contribute to the force resultants N_{jk} further terms of the order h^3, while not affecting the values of M_{jk}. It appears, however, that nothing is gained by including these further terms since their correctness is not assured due to the fact that terms proportional to ζ^2 have already been disregarded the expression for the displacement vector. The advantage of retaining term of order h^3 in the N_{jk} although they are not likely to be correct (those obtaine from (4) or those obtained by developing ϵ and γ one step further) lies in th fact that thereby the equation of moment equilibrium about the normal to th shell reduces to an identity as it should in view of its being obtained b integrating over the shell thickness the identity $\tau_{12} = \tau_{21}$. If terms of orde h^3 are not retained in the N_{jk} this moment equation is not quite satisfied excep in the case of a spherical shell or flat plate. In actual calculations these h terms are usually not included and the slightly violated moment equation disregarded. The developments in [2] and [1] represent an exception in this respect. There, for a circular cylindrical shell, the correction terms whic would follow from a development of the strains (1) and (2) up to ζ^2 a derived by direct geometrical considerations for the circular cylindrical shel

---

[3] See equations (20) to (24) in [5].

It is important to notice that in the M_{jk} only terms of the same order in h occur and it seems not justified to disregard generally any of them.

As a special example of the general formulas (9) to (16) the case of a cylindrical shell of any cross section may be considered in detail. Taking as coördinates on the middle surface the axial length $\xi_1 = x$ and the circumferential arc length $\xi_2 = s$ and writing

$$u_1 = u, \qquad u_2 = v, \qquad \frac{\partial}{\partial x} \equiv (\ )', \qquad \frac{\partial}{\partial s} \equiv (\ )^{\cdot}$$

we have

$$\alpha_1 = \alpha_2 = 1, \qquad R_1 = \infty, \qquad R_2 = R(s). \tag{17}$$

From (4) in conjunction with (1) and (2) follow

$$\epsilon_x{}^0 = u', \qquad \epsilon_s{}^0 = v^{\cdot} + w/R, \qquad \gamma_{xs}{}^0 = v' + u^{\cdot}. \tag{18}$$

The changes of slope, as defined in [5], Eq. (32), are

$$u_x' = -w', \qquad u_s' = -w^{\cdot} + v/R, \tag{19}$$

giving in accordance with Eqs. (43) and (44) in [5]

$$\kappa_x = -w'', \qquad \kappa_s = -(w^{\cdot} - v/R)^{\cdot}, \qquad \tau = -2w'^{\cdot} + (v/R)' \tag{20}$$

so that, according to (5), (6) and (7)

$$\left\{\begin{aligned} \kappa_x{}^* &= -w'', \qquad \kappa_s{}^* = -w^{\cdot\cdot} - w/R^2 + v(1/R)^{\cdot} \\ \tau^* &= -2w'^{\cdot} + (v' - u^{\cdot})/R \end{aligned}\right. \tag{21}$$

Since $1/R_1 = \infty$ M_{ss} and M_{sx} are directly given in terms of (21) while the combinations of (18) and (21) occurring in M_{xx} and in M_{xs} have the values

$$\left\{\begin{aligned} \kappa_x{}^* + \epsilon_x{}^0/R &= -w'' + u'/R \\ \kappa_s{}^* + \epsilon_s{}^0/R &= -w^{\cdot\cdot} + (v/R)^{\cdot}. \\ \tau^* + \gamma_{xs}{}^0/R &= -2w'^{\cdot} + 2v'/R \end{aligned}\right. \tag{22}$$

For a circular cylindrical shell these expressions are further simplified, $R(s)$ becoming the constant cross-sectional radius a.

The results obtained for the moment resultants of the circular shell are listed, in comparison with the expressions used by other writers, in the following table

	[1], [2] and here	[3], [6], [7]	[4]	Simplified Expressions
$\frac{(1-\nu^2)}{Eh^3} M_{xx}$	$w'' - u'/a$ $+ \nu(w^{\cdot\cdot} - v^{\cdot}/a)$	w'' $+ \nu(w^{\cdot\cdot} - v^{\cdot}/a)$	w'' $+ \nu(w^{\cdot\cdot} + w/a^2)$	w'' $+ \nu w^{\cdot\cdot}$
$\frac{}{3} M_{xs}$	$2w'^{\cdot} - 2v'/a$	$2w'^{\cdot} - 2v'/a$	$2w'^{\cdot} - 2v'/a$	$2w'^{\cdot}$
$\frac{}{3} M_{sx}$	$2w'^{\cdot} - (v' - u^{\cdot})/a$	$2w'^{\cdot} - 2v'/a$	$2w'^{\cdot} - 2v'/a$	$2w'^{\cdot}$
$\frac{(1-\nu^2)}{Eh^3} M_{ss}$	$w^{\cdot\cdot} + w/a^2$ $+ \nu w''$	$w^{\cdot\cdot} - v^{\cdot}/a$ $+ \nu w''$	$w^{\cdot\cdot} + w/a^2$ $+ \nu w''$	$w^{\cdot\cdot}$ $+ \nu w''$

It is seen that general agreement exists for M_{xs} but that different versions occur for the other moment resultants. The reason for this is that in some instances κ has been used for κ^* [3], [6], [7], and that terms containing ϵ^0/a have been neglected [3], [4], [6], [7]. It is noticed that although these terms have been neglected in [3], [4], [6], [7] and although neither τ nor $\tau + \gamma^0/a$ nor τ^* are identical with $\tau^* + \gamma^0/a$ the value of M_{xs} is still given correctly.

From the form of the relations which are tabulated it is not apparent that the differences between them are always of negligible influence on the results of calculations. It appears, however, that when they are of negligible influence, the further simplified relations as given in column 4 of the table, which ar identical with those of the theory of flat plates, should be equally justified.[4]

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES

1. C. B. Biezeno, R. Grammel, *Technische Dynamik*, Berlin (1939).
2. W. Fluegge, *Statik und Dynamik der Schalen*, Berlin (1934).
3. A. E. H. Love, *Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambrid (1927).
4. J. Prescott, *Applied Elasticity*, London (1924).
5. E. Reissner, "A new derivation of the equations for the deformations of elasti shells," *American Journal of Mathematics*, vol. 63 (1941), pp. 177-184.
6. S. Timoshenko, *Theory of Elastic Stability*, New York (1936).
7. S. Timoshenko, *Theory of Plates and Shells*, New York (1940).

UNIVERSITY ✻ CALCUTTA
LIBRARY
CENTRAL

---

[4] It may be noted that in the case of a spherical shell ($R_1 = R_2 = a$) there is agreement between the formulas (13) to (16), Love's abbreviated expressions and those contained in [5].